THE AMERICAN
JOURNAL OF PSYCHOLOGY
Founded in 1887 by G. STANLEY HALL

Vol. LI JULY, 1938 No. 3

# A THEORETICAL FIELD-ANALYSIS OF AUTOMOBILE-DRIVING

BY JAMES J. GIBSON, Smith College, and LAURENCE E. CROOKS,
Society of Automotive Engineers

Of all the skills demanded by contemporary civilization, the one of driving an automobile is certainly the most important to the individual, in the sense at least that a defect in it is the greatest threat to his life.[1] But despite the consequent importance of knowledge about the nature and acquisition of this skill, no more than a beginning in this direction has been made by psychologists, and that chiefly in the field of devising tests to measure some of its inferred components.[2] A systematic set of concepts is needed in terms of which we can describe precisely what goes on when a man drives an automobile, and such a theory, if it is to be useful, must have practical as well as psychological validity. The following paper has been written in the effort to make a systematic description of this sort.

When this undertaking was first proposed, the effort was made to base the analysis upon the more familiar concepts of present-day psychology—habits, attitudes, and response-sequences. In this effort, however, the writers had but small success. Very little in the way of a useful theory emerged. They finally concluded that the task of the automobile driver is so predominantly a perceptual task, and that the *overt* reactions are so relatively simple and easily learned, that the analysis has to be carried out on a perceptual level and with concepts more appropriate to this requirement.

* Accepted for publication August 1, 1937. The authors state: "This article is the outcome of discussions between a psychologist and a practical student of driving."

[1] Accident statistics are now widely publicized. Cf., *e.g.*, R. B. Stoeckel, M. A. May, and R. S. Kirby, *Sense and Safety on the Road*, 1936, 211, 242. Over a period of a year, one driver in every twenty is involved in an accident resulting in death or injury.

[2] See, for example, H. R. De Silva, *Research in Driving Skill*, Mass. State College, 1935.

Their alternative was, therefore, to utilize concepts like the 'field' of the driver, 'valences,' and the general cross-sectional method employed by Lewin.[8]

From this starting point it was possible to formulate a theory better suited to the facts at hand. We may assume that driving is a type of locomotion through a 'terrain' or field of space. It is psychologically analogous to walking or running, except that driving is locomotion by means of a tool; that is to say, the automobile can be thought of as a tool for more effective locomotion in the same way that a club is a tool for more effective striking. The primitive function of locomotion is to move the individual from one point of space to another, the 'destination.' In the course of simple locomotion 'obstacles' are met with, *i.e.* are seen, and the locomotion itself must be appropriately modified in order to avoid them. 'Collision' with an obstacle stops locomotion and may produce bodily injury. Locomotion is therefore guided chiefly by vision, and this guidance is given in terms of a 'path' within the visual field of the individual such that obstacles are avoided and the destination ultimately reached. These concepts of *terrain, destination, obstacle, collision* and *path* should be applicable to any type of locomotion, whether that of the infant learning to walk, the open field runner in football, or the operator of an automobile. It might seem more logical to apply them first to the simpler types, but nevertheless we shall attempt to work them into a theory of automobile driving, with the idea that a general hypothesis which would cover the more difficult problem would assuredly include valid principles for the solution of the simpler ones.

## The Field of Safe Travel and the Nature of Steering

The visual field of the driver is a rather special sort of field in several respects. It is selective in that the elements of the field which are pertinent to locomotion stand out, are attended to, while non-pertinent elements, such as 'scenery,' normally recede into the background. The most important part of the terrain included in this pertinent field is the road. Within the boundaries of the road lies, according to our hypothesis, an indefinitely bounded field which we will name the *field of safe travel*. It consists, at any given moment, of the *field of possible paths which the car may take unimpeded*. Phenomenally it is a sort of tongue protruding forward along the road. Its boundaries are chiefly determined by objects or features of the terrain with a negative 'valence' in perception—in other

[8] Kurt Lewin, *Topological Psychology*, 1936, 1-

words obstacles. The field of safe travel itself has a positive 'valence,' more especially along its mid-line. By *valences,* positive or negative, we refer to the meanings of objects by virtue of which we move toward some of them and away from others. The valences of objects with respect to locomotion may be quite different ones from their valences with respect to eating or esthetic enjoyment when the individual is not simply propelling himself between them. For instance, a hot-dog wagon has a negative valence with respect to locomotion, but a positive one with respect to appetite.

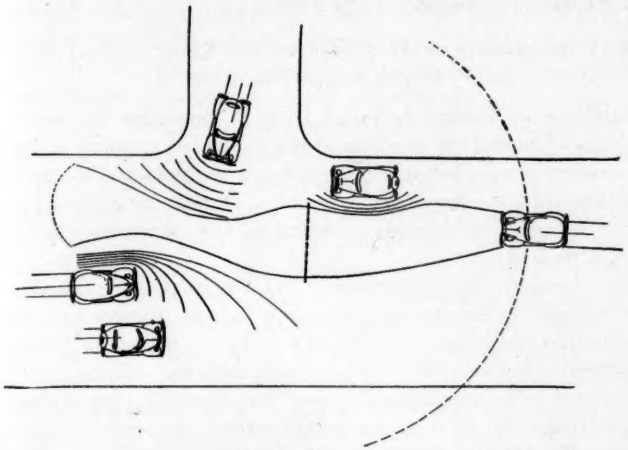

FIG. 1. THE FIELD OF SAFE TRAVEL AND THE MINIMUM STOPPING ZONE OF A DRIVER IN TRAFFIC

(If, in this and the following figures, the page is turned around and the figure is viewed from what is now the right, the reader may the better be able to empathize the situation, since he will then have the point of view of the driver of the car whose field of safe travel is under discussion.)

Fig. 1 is a representation of this field at a specific instant. The field of safe travel, it should be noted, is a spatial field but it is not fixed in physical space. The car is moving and the field moves with the car *through* space. Its point of reference is not the stationary objects of the environment, but the driver himself. It is not, however, merely a subjective experience of the driver. It exists objectively as the actual field within which the car can safely operate, whether or not the driver is aware of it. It shifts and changes continually, bending and twisting with the road, and also elongating or contracting, widening or narrowing, according as ob-

stacles encroach upon it and limit its boundaries. As will later be evident, other factors may also serve to constrict these boundaries, but for the present we may point out that its depth is limited by frontal obstacles such as other vehicles, policemen and stop lights, and its width by 'marginal' obstacles like curbs, ditches, soft shoulders, walls, parked cars, pedestrians or 'white lines.'

It is now possible to define precisely the operation of steering an automobile. *Steering, according to this hypothesis, is a perceptually governed series of reactions by the driver of such a sort as to keep the car headed into the middle of the field of safe travel.*

## The Determinants of Accelerating and Decelerating, and the Minimum Stopping Zone

Collisions are avoided during all kinds of locomotion by one of two methods—changing the direction of the motion or stopping it. For the driver of the car this means steering away, or decelerating. Having offered a hypothesis for the former, we have next to consider the problem of what factors determine the response of taking the foot off the throttle or stepping on the brake.

First, however, the question arises: What initiates and maintains locomotion itself? In the simplest case we may suppose that speed of locomotion is a function of the urge within the individual toward his destination. This directed motive—'vector' if one prefers the word—may be consciously represented by the experience of 'hurry.' It is, of course, perfectly true that people frequently drive an automobile without any specific destination. Driving 'for pleasure' may be considered, however, in common with most play, the using of a tool or a skill for its own sake. If driving is locomotion with a tool—a very ingenious, elaborate, and versatile tool—it is not surprising that like most tools its manipulation, once acquired, may be an end in itself. Speed of locomotion in this situation becomes a very complex matter; but, limiting ourselves to the simpler case, it appears that locomotion is maintained and augmented by a motive of the simplest kind to be found in all behavior—that of 'going somewhere.'

Accelerating is then a function of this motive. To what may we ascribe decelerating? Once more, to take the simplest case and disregarding such limits on speed as legal taboos, we may suppose that it occurs when the field of safe travel visibly contracts, and that stopping, like steering, is primarily an avoidance reaction to obstacles. In the diagram of Fig. 1, for example, if the car ahead, which is apparently passing the other, is instead

turning into the side street, then the field of safe travel will be cut off and our driver will, it is to be hoped, stop.

We do not, however, have as yet a sufficiently precise description. How much contraction of the field of safe travel can occur, and what is the relation between contraction and deceleration? The answer is given by another field concept. There is within the field of safe travel another zone, phenomenally less precise but behaviorally and objectively just as real, which is set by the minimum braking distance required to stop the car. It is the zone within which our driver could stop if he had to, and it can be supposed to be present implicitly in every driver's field. How accurately it accords with reality is another question. Unlike the field of safe travel, the size of this *minimum stopping zone* is dependent on the speed of the car—and also, it may be noted, on the condition of the road-surface and of the brakes.[4] The driver's awareness of how fast he is going does not consist of any estimate in miles per hour; instead, he is aware, among other things, of this distance within which he could stop.

Normally, the forward margin of the minimum stopping zone is well behind the forward boundary of the field of safe travel. Speed may be increased up to the point where the zone nears the size of the field, and when this happens driving begins to feel 'dangerous.' For the experienced driver, the car is in a sense projected to its potential location at the front of this zone. Hence, when the field of safe travel contracts by reason of an abstacle which encroaches upon it, deceleration occurs *in proportion as the forward margin of the field recedes toward the minimum stopping zone.*

A sudden frontal contraction or a shearing off of the field of safe travel, which cuts it down to or below the minimum stopping zone, produces in the driver a feeling of imminent collision, sometimes approaching panic, and an immediate and maximum braking reaction. There is an 'emergency.' Much more frequently, however, there occurs a gradual contraction of the field, and, as it approaches the front boundary of the zone, there follows a gradual slowing reaction of such strength as to keep the zone continually smaller than the field.

The ratio of depth-of-field to depth-of-zone which tends to be maintained in given traffic conditions by a given driver is probably to a large extent habitual. It might be designated as the *field-zone ratio* and thought

[4] Tables giving the minimum stopping distance for different speeds under standard conditions have been widely publicized. See, for example, Stoeckel, May, and Kirby, *op. cit.*, 1936, 176.

of as an index of cautiousness. The ratio may be expected to decrease when the driver is in a hurry.[5]

Expressions of the foregoing principle may be found in such commonsense rules as never driving so fast that one cannot stop within the 'assured clear course,' and never 'overdriving' one's brakes or one's headlights. Inattentiveness in the driver usually means that objects in the terrain or inside the car which are not pertinent to locomotion stand out in his visual field and that consequently his field of safe travel, if it exists at all, may become incorrectly bounded. The *objective* field of safe travel contracts until it threatens to get below the stopping-zone, or does so, without a corresponding contraction of the *behavioral* field. Frequently, it may be suggested, the remedy is not merely increased vigilance or tension in the driver, but the development by learning of semi-automatic perceptual habits and motor habits such that a safe margin is maintained between stopping zone and field of safe travel.[6]

When an obstacle suddenly cuts off the field of safe travel well inside of the stopping zone—an 'emergency'—an entirely new field may open up which did not exist an instant before. Obstacles such as the shoulder of the road, or the curb, or even a shallow ditch which just previously had a minus valence may take on a positive valence and become no longer obstacles but a field of travel. The possibilities for the opening up of a new field depend on the driver's sensitivity to the pertinent field as a whole, and upon the degree to which this field is imbued with meaning for locomotion. The negative valences of obstacles vary in strength, some being more-to-be-avoided than others. The experienced and skillful driver is able to shift his field of travel with considerable flexibility in case of an 'emergency' and thereby avoid a collision. In such a case the field of safe travel is highly relative. It might be called the field of safest travel.

---

[5] Except for emergencies, more efficient brakes on an automobile will not in themselves make driving the automobile any safer. Better brakes will reduce the absolute size of the minimum stopping zone, it is true, but the driver soon learns this new zone and, since it is his field-zone ratio which remains constant, he allows only the same relative margin between field and zone as before.

[6] The relative importance of effortful attention and of habit in safe driving needs to be worked out. It seems to be assumed, on the basis of what motorists say to police officers after accidents, that most accidents are 'caused' by inattention. Hence, it is concluded, drivers need to be warned sternly of their responsibility to be attentive. Such an inference is characteristic of the theological psychology of legal thinking. It is more likely that drivers need to be taught safe habits, among which is the habit of attending to the field-zone ratio. Contemporary safety campaigns tend to emphasize the factor of vigilance and to play upon the fear-motive to such an extent that the development of correct habits may be hampered in some drivers. These campaigns (*e.g. And Sudden Death*) may be more successful in producing attitudes of timidity than in improving the average driver's performance. If drivers *knew how* to drive safely most of them would do so.

The bearing of our hypothesis on the problem of high speed driving is not hard to make out. A good road which is free from traffic offers a very much expanded field of safe travel. The driver's urge toward his destination, especially if he is one who tends to maintain a fairly constant field-zone ratio, will result in high speed. He tends to feel that a higher ratio between field and zone than is demanded by safety is wasteful of time. With increasing speed, however, the minimum stopping zone extends farther and farther down the road. Accordingly the field of safe travel must at all times extend even farther down the road in proportion. Even with a perfectly clear highway a limit to this extension is eventually reached, since obstacles become harder to see as their distance increases, the acuteness of human vision being what it is. Since, at very high speeds, the driver's attention must be fixed at greater

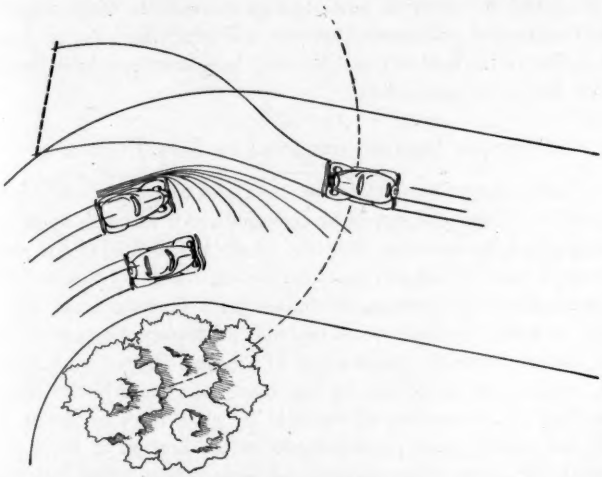

FIG. 2. SUDDEN REORGANIZATION OF THE FIELD OF SAFE TRAVEL TO MEET AN EMERGENCY

An unforeseen contraction of the old field below the minimum stopping zone.

and greater distances—for he has to operate farther and farther ahead in his visual field and scan the road for increasingly minimal cues—the purely sensory task of seeing the zone and the field correctly is more difficult, the strain is greater, and chances for error multiply.

These two fields which have been proposed cannot of course be conceived of as visible, strictly speaking, in the sense than an object with a contour is visible. Nor are their boundaries sharply defined as are lines and contours, although for convenience we may diagram them as if they were. They are *fields within which certain behavior is possible*. When they are

perceived as such by the driver, the seeing and the doing are merged in the same experience. Since their boundaries are set by behavioral possibilities, these boundaries are more or less correct and appear as transitions more or less sharp depending on the skill and experience of the driver. Driving skill largely consists, we suggest, in the organization, within the pertinent visual field of the driver, of a correctly bounded stopping zone for the entire repertory of speeds, roads, and surface conditions, and a field of safe travel which is precisely moulded to the actual and potential obstacles in the total field at any given instant. In other words, the field must be shaped in accordance with the objective possibilities for locomotion. If we add that there are concurrent motor reactions which maintain a safe and constant ratio between these two, and which keep the car running in the middle of the field of travel, we shall have described the basic operations of driving an automobile.

## Factors Limiting the Field of Safe Travel

*(1) Natural boundaries.* So far we have said that the driver's field of safe travel is limited and shaped by obstacles which encroach upon it, but this statement is by no means complete. Limits to this field of the possible seen paths which the car may take may be set, not only by obstacles, but also by physical and physiological factors. Such limitations are, first, the distance at which daylight vision becomes inadequate, because of purely optical factors; secondly, the margins of the field lighted by headlights; thirdly, constriction of the field by fog, snow, *etc.*; fourthly, the diminishing or even the destruction of the field by glare from oncoming headlights; and, finally, such physical limits as the horizon at the top of a hill which the driver is approaching. All these factors, when they cause a diminishing of the field-zone ratio, should bring about a reaction of deceleration. Outside of these natural boundaries of the field there lurk what we shall call *potential* obstacles.

*(2) Inflexibility at higher speeds.* Another fact which limits the field of potential paths of the car is that it may not include sharp turns at high speed. The field, of course, conforms to the turns of the road and to obstacles on its edges, but it cannot conform by curving more than a certain amount for a certain speed without the occurrence of skidding. One may conceive that the centrifugal force which produces a skid operates as a kind of potential obstacle which encroaches on the concave side of a field of safe travel (Fig. 3). The field becomes less flexible, more rigid or straighter, with higher speed, for the reason that, if one turns too sharply,

one is literally likely to 'run into' a skid. This description of skidding may at first sound implausible because we tend to identify skidding with the obstacle *into which* we may skid; but an examination of Fig. 3 will show that, while the field of possible paths is bounded on the right by actual obstacles, it is bounded on the left by the potential skid which would occur if the car were swung into it. At a given speed, the feeling of losing traction with the road is projected into the terrain ahead in such a way as to make certain paths impossible.

*(3) Obstacles and their 'clearance lines.'* It has already been stated that obstacles limit the field of safe travel by virtue of their minus valence for

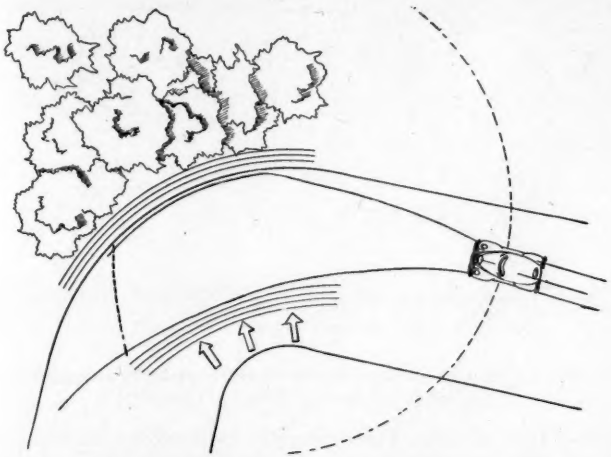

FIG. 3. INFLEXIBILITY OF THE FIELD OF SAFE TRAVEL AT HIGH SPEED
The field is bounded on the right by fixed obstacles and on the left, at this speed, by a projected feeling of the potential skid.

locomotion. The open terrain between the obstacles, if it leads off toward the destination, has a plus valence. This negative valence of an obstacle, we could suppose, is an indicator of the consequences of collision with the object in question. The child learning to walk about probably learns these valences rather readily. An object or place which is attractive to the child induces locomotion toward it. In the course of such locomotions, the child bumps against fixed objects in his field, such as furniture, and, as a result of these collisions, acquires a repertory of avoidance responses to them. These avoidances are superimposed on the general activity of locomotion,

and, taking place as they do only in the immediate neighborhood of the obstacle, they can be represented in the child's field by a negative valence of the obstacle and of the space immediately surrounding it. We can thus suppose that every obstacle comes to have a sort of halo of avoidance, and that this halo can be represented by 'lines of clearance' surrounding it. In analogy with contour lines on a map, we could draw lines of equal negative valence around an obstacle, each line as it gets closer to the obstacle representing a greater intensity. Locomotion in a field containing obstacles would then be a matter of moving *toward* the destination and at the same

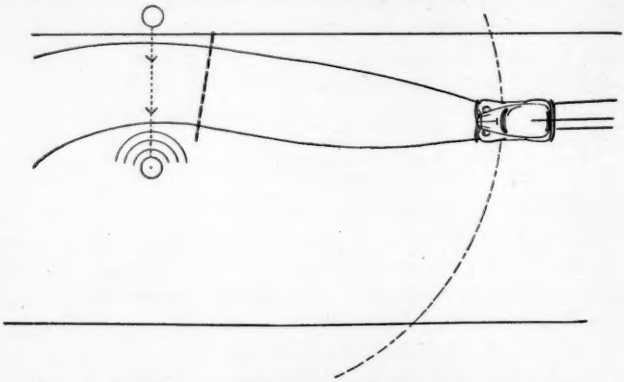

FIG. 4. THE FIELD OF SAFE TRAVEL AS DETERMINED BY THE *future* CLEARANCE-LINES OF A MOVING PEDESTRIAN

time *away* from obstacles. The resultant of these two tendencies, summating the motives (vectors) at each point, yields the *path* of locomotion.[7]

For the automobile driver these concepts can be taken over in their entirety. The driver has learned at least the essentials of locomotion when he learned to walk, and learning to drive is mostly learning the use of the tool. His field of safe travel is at every moment bounded and shaped by the halos of obstacles in the terrain. It is a reasonable hypothesis, moreover, that the more injurious a collision with an obstacle would be, the more it is avoided and the greater is the extent of its clearance-lines. Some

[7] The possibility that this analysis carried further is capable of explaining 'roundabout' behavior is worth considering. Lewin suggests that the detour performance requires a "restructuring" of the field and "insight"—an extra hypothesis which may not be necessary.

obstacles would then affect the field of safe travel more than others. For example, a heavy bus approaching at great speed would have more clearance-lines than a model-T Ford; and a precipice would have more clearance-lines than a shallow ditch.

*(4) Moving obstacles.* We are now in a position to deal with obstacles more fully. Two cases must be distinguished: that for stationary and that for moving obstacles. In the former case (curbs, trees, parked cars) the clearance lines radiate from the physical location of the obstacle and affect the field of safe travel as it impinges on this location. Figs. 1-3 have repre-

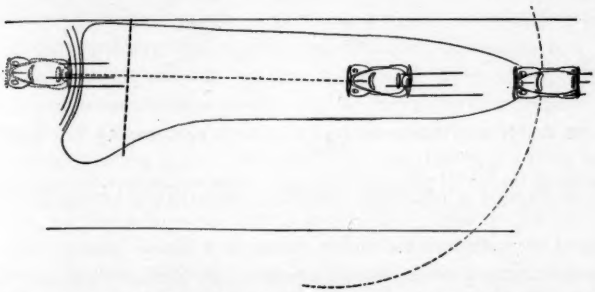

FIG. 5. THE FIELD OF SAFE TRAVEL AT RELATIVELY HIGH SPEED, WHERE A MOVING OBSTACLE IS PHYSICALLY WITHIN THE FIELD BUT WHERE ITS PROJECTED CLEARANCE-LINES DETERMINE THE FORWARD MARGIN OF THE FIELD

sented this state of affairs. For a moving obstacle (vehicle, person, animal) the clearance lines radiate from *the point where the obstacle will be when the driver's car comes closest to it,* or in other words from the point of potential collision. Consequently one may for example steer *toward* the present position of a moving pedestrian (Fig. 4), since, in effect, one is steering away from where he will be when the car reaches him. Likewise a car ahead going in the same direction may, at high speeds, actually be inside the field of safe travel without affecting it, since its clearance-lines are projected far ahead in the field where it would be if it stopped (Fig. 5). In general, the greater the speed of a moving obstacle, the farther ahead of it are its clearance-lines projected. Likewise, for all obstacles moving toward the driver, whether from the front or from one side, the greater their speed, the more do their clearance lines encroach upon the field of safe travel. Finally, the more unpredictable the movement of an obstacle, the more extensive are its clearance-lines. Examples would be

cars behaving 'strangely,' blind men, children (especially on bicycles), and old people. A street-car should have few and narrow clearance-lines to the side, since its lateral movement is usually predictable and negligible.

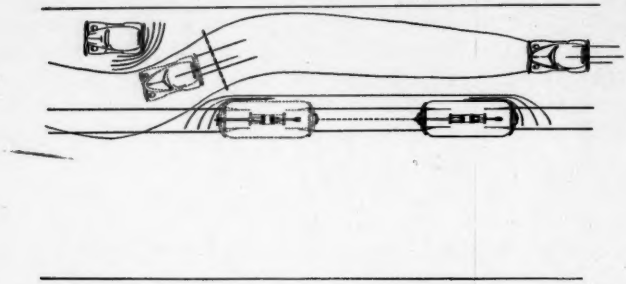

FIG. 6. A DRIVER ATTEMPTING TO PASS A STREET-CAR BEFORE REACHING A PARKED CAR

At this instant the field of safe travel looks open, since the clearance-lines of the trolley are projected as indicated for the moment when he will be passing it.

A parked car with a driver visibly sitting in it should have a wider halo than one without a driver, since the former may suddenly pull away from the curb.

It should be remarked here that the correct location of the clearance-lines of a moving vehicle ahead demands, not only a sort of projection of

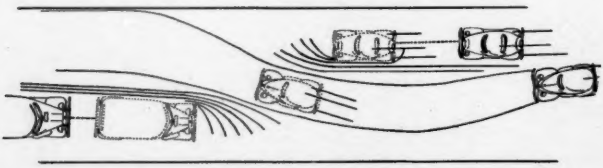

FIG. 7. A DRIVER PASSING WITH INTENT TO CUT IN

An integration of three perceived speeds of movement is involved in the seeing of this open field of safe travel.

this vehicle to the proper point in its own path, but also the correct projection of one's own car to the point of intersection of the two paths. Both one's own and the other person's field of safe travel are jointly involved. An example is given in Fig. 6. In other words an 'estimate' of the speed of the obstacle plus an 'estimate' of the speed of one's own car are com-

bined—although the term *estimate* is misleading since there seems to be no conscious process of calculation involved. Here is a case of a highly complex situation, involving relationships between two speeds of movement, which would not be an easy problem to solve with pencil, paper, and formulae; but for the skillful driver the perceptual field-situation may be immediate, clear, and (let us hope) accurate. Complex 'estimates' of speed and location are represented in the experience of the driver only by the simple seeing of an open or a closed field of safe travel. In Fig. 7, which represents a driver passing the car ahead in the face of an oncoming driver, there is involved a projection of *three* cars, with *three* speeds of movement, to their correct positions at a future instant. Here indeed is a very extraordinary type of perceptual 'constancy-phenomenon.'[8]

*(5) Potential obstacles.* In addition to the actual obstacles which at any moment are encroaching upon the field of travel to some degree, there may be in the pertinent field *barriers to sight* which also throw out lines of clearance. For the trained driver, any intervening object from behind which a moving obstacle might suddenly cut into the field of travel—a building, a blind corner, or even a parked car—radiates clearance lines (Fig. 8), because it has a secondary or indirect negative valence. The limits to the field of safe travel imposed by darkness beyond the headlight-zone, by fog, and by curves or by the brow of a hill ahead, likewise constitute such barriers. Behind them there are *potential* obstacles. There may or may not be *actual* obstacles behind them; the objective field of safe travel may or may not be clear; but, since the driver can react only to his subjective field, the latter is properly shaped and limited by these barriers.

It must be admitted that for the average driver potential obstacles do not have the potency of visible obstacles. Very frequently one of these barriers to sight is not recognized, its negative valence is not perceived, and the field of safe travel is incorrectly related to the objective possibilities for locomotion. Psychologically this fact is not surprising, for the existence of a potential obstacle behind the barrier must be learned by experience and its negative valence transferred to the barrier. Probably each new type of sight-barrier which is met necessitates a new process of learning. More frequently, a barrier representing potential obstacles of

[8] The movement-perceptions of the three objects in question are subject in a very striking way to the principle of 'constancy.' When one considers the actual stimulus-cues for the perceived speed of the approaching car, the to-be-passed car, and finally the driver's own car, and realizes that these speeds are perceived in relation to a phenomenally stationary road, the degree of 'phenomenal regression' away from the stimuli and toward 'reality' is astounding.

low valence is disregarded. Even the most cautious driver undoubtedly drives on faith under certain circumstances. For example, a blind stop-street which enters the highway ahead does not usually affect the field of safe travel, although occasionally it conceals an automobile which will not stop at the intersection. A green traffic-signal is interpreted as a *guarantee* that the field will remain open and that unseen cars on the crossroad will not enter it. It is invariably assumed that the road does not end abruptly

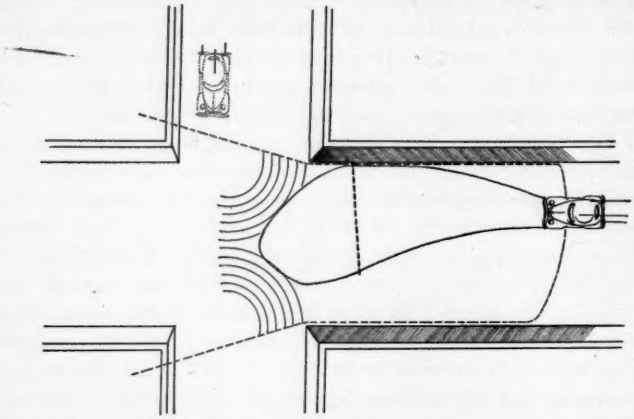

FIG. 8. A BLIND CORNER CONSTITUTING A BARRIER TO VISION AND ITS EFFECT ON THE FIELD OF SAFE TRAVEL

At this moment the clearance-lines of potential obstacles cut off the field.

in a stone wall just out of sight over the brow of the hill ahead—a very probable assumption, to be sure, but still an assumption. It is also generally assumed that the absence of warning signs by the highway means the absence of obstacles for which one should be warned.

The driver's field of safe travel, as he reacts to it, is delimited, we now see, by *probable* as well as by seen obstacles. Equally, it is not delimited by *improbable* obstacles. The more probable types of obstacles are usually indicated by highway signs, crossing signals, and the like, to which the driver supposedly learns to react as he would to the obstacles themselves. But the highway signs vary considerably in different parts of the country, and they are usually put up without any regard for the actual probabilities of the obstacles in question. The average driver becomes dimly aware that he must calculate the probabilities of the road for himself without much dependence on highway signs and consequently comes to perceive them as merely legal obstacles to which he must conform, when he does, for reasons distinct from the necessities of locomotion. Their negative valence is based on potential legal consequences rather than potential collision.

*(6) Legal obstacles and legal taboos.* The field of safe travel is modified, not only by the various types of obstacles already described, but also by perceived symbols—more specifically by signals, gestures, signs and markings. The traffic signal showing red and the traffic policeman with arms outstretched are the most obvious examples. Arm signals and the stop-lights of the car ahead are cues which enlarge its clearance lines or shift them to one side. Highway markings indicating lanes and the areas around fire-plugs are symbolic obstacles, as are "private" signs on driveways and roads. Finally, "stop" signs at intersections, blinkers, and a variety of highway signs, "slow," "curve," "school," "speed limit 20 m.p.h" and the like, affect the field of safe travel with varying degrees of effectiveness. Too frequently they are not reliable indicators of the objective possibilities for locomotion and, as we have suggested, become purely legal taboos. A sign which is not occasionally 'reënforced' will in the end no longer stand for anything, as students of the conditioned response are well aware.

There are, accordingly, in the pertinent field of the driver, obstacles whose negative valence is partially or wholly legal. This fact is illustrated by their greatly increased negative valence when a traffic policeman is present in the field as compared with the same field without the policeman. Likewise, manoeuvres, which are warranted by the visible field-conditions but are legally tabooed, are seldom made in the visible presence of the law, as for example passing on curves where the field of view is unrestricted, or at a clear intersection where another road enters only from the right. Traffic laws being absolute can agree only roughly with the actual facts of dangerous driving. From the point of view of the driver who undertakes to drive safely but also skillfully and efficiently, the unrealistic legal restrictions present a problem of compromise.

## The Field of the Other Driver

It has been said that driving an automobile is a highly individualistic activity. There is no way of communicating with another driver except by way of the horn, or by arm signals—neither of which is a particularly sociable mode of interaction. As a consequence perhaps of this anonymity on the road, the situation is one conducive to competitive rather than coöperative attitudes. Each driver tends to think of his own task and to regard the other drivers merely as obstacles.

On the other hand, the driver does take into account the field of the other person. It is always assumed, for example, that *the other driver sees you.* The clearance lines of an approaching car are projected, not as if it were simply an inanimate moving obstacle or an incalculable agent, but

on the assumption that it is avoiding you as much as you are avoiding it. More specifically, it is assumed that the other driver's field of safe travel does not overlap yours. To speak of an 'assumption' is, of course, inadequate, but the behavioral fact remains. For example, when a driver sees two cars approaching on a two-lane road and the rear car swings out to pass with the intention of cutting in before the various paths meet, the driver is usually very much aware of his own menacing aspect in the field of the passing driver. Frequently he uses his advantage to the utmost,

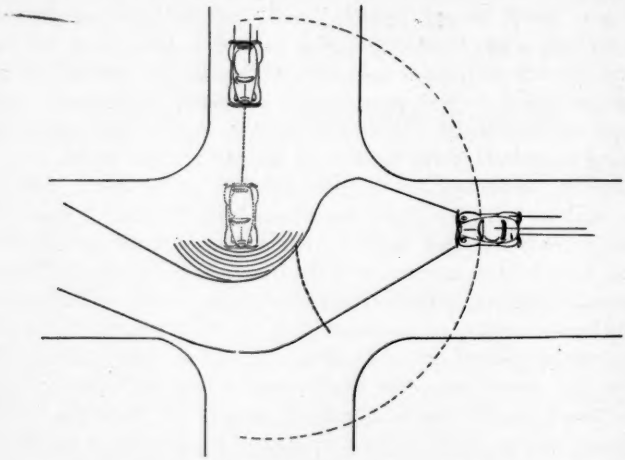

FIG. 9. BLUFFING AT AN INTERSECTION

A driver projecting his field of safe travel on the assumption that he is a formidable obstacle in the other driver's field.

maintaining speed despite his dubiously open field of safe travel, in the bold manner of a dog facing a canine trespasser on home ground.

Another example is the use of the horn, by which one announces one's presence as an obstacle encroaching on the other person's field. If the other person behaves as if he did *not* see you—if his field of safe travel is apparently incorrect—then his clearance lines encroach unduly on one's own field. In this situation blowing the horn simply expands one's *own* clearance lines in the other person's field, advertising that "I am a dangerous obstacle," and one goes on the faith that this expansion of clearance-lines in the other's field will indirectly clear up one's own at present constricted field. The obstacle, one trusts, will stop.

Occasionally, however, the obstacle itself may decide to advertise its dangerousness and behave in a powerful and unyielding fashion. We then have the situation of mutual bluffing. Each driver expands his own clearance lines to the utmost, maintaining speed or using the horn, and they may persist even up to the point where their respective minimum stopping zones touch. At this point we may leave them to their fate.

## The Car Itself as a Field

The perceptions of the driver are organized into a number of spatial fields, each being a sort of projection of the behavioral opportunities, each being a space within which certain behavior is possible and outside of which it is impossible. Within the field pertinent to locomotion are the field of safe travel, the minimum stopping zone, and on some occasions the field of the approaching driver or pedestrian. When the driver prepares to stop, overtake, or make a turn, his pertinent field *includes the road behind* for the reason that he must not encroach too suddenly on the field of safe travel of a potential driver who is coming up fast from the rear. The rear-vision mirror is the usual mediator of this rearward field. Finally, there is still another field which needs to be taken into account—that of the car in which the driver is sitting. We suggested that the car was a tool whose use was locomotion. It is also a sort of field which yields a variety of perceptual cues and which invites and supports specific actions. The impressions constituting it are kinesthetic, tactual, and auditory, as well as visual, and they interact with the impressions from the terrain to produce the totality of cues on which the driving-process is based. The 'feel' of the car or the 'behavior' of the car are terms which indicate what is meant by this particular field of experience.

In the first place, locomotion involves a feeling of contact with the terrain together with a feeling of exerting force against it. In the case of the driver these feelings are referred to the car; the actual impressions are secondary and indirect. The feeling of driving on an icy surface is similar to the feeling of walking on ice despite the different cues and different muscular movement involved. The perception of *traction* is therefore an important feature of the driving process. Under extreme conditions, as when rain freezes on the road, locomotion takes place predominantly by means of kinesthetic and tactual cues from the car-field. The driver 'feels his way' along the road. It is evident that he projects his actual impressions into the car-field and may be said to feel the road with the car in the same sense that one feels the ground with a walking-stick.

In the second place locomotion involves a constant orientation with respect to the path. The car, as we have already suggested, is kept pointed into the middle of the field of safe travel; the two are kept in alignment. The car is, therefore, something which can be oriented, pointed or aligned, and this characteristic depends on those sensory cues which go to make up the car-field. During a skid the car becomes disoriented with respect to its path. If the driver, however, maintains *his own* orientation to it and if the skid is only a moderate one, he can correct the car's orientation by simply steering into the middle of the field of safe travel.[9] But if the skid turns the car completely around or if the driver is unskilled, his own personal orientation to the path may be lost. The car is then 'out of control,' primarily because the *driver* is disoriented.

In the third place, locomotion involves the maintenance of an upright position. The tilt of the car on a high-crowned road and the corresponding pull of the wheel, the centrifugal force and the sway on a sharp curve, all contribute to the driver's sense of his car and the feeling of its stability or lack of it.

Enough has been said to indicate the nature of the car field. This field which the driver inhabits is one in which the impression and the action are especially intimately merged. The impression and the action imply one another. In common with other tools, it has a set of potentialities—a certain repertory of performances which the skillful driver knows and utilizes fully and of which even the tyro is forced to take some account. A high degree of driving skill demands a complete sensitivity to and control over this field. If this demand is fulfilled, the car tends to become, like any properly used tool, simply a sort of physical extension of the driver's body.

## Application

The situation which faces highway-safety officials in dealing with a public who insist on killing themselves at an ever-increasing rate is similar to that of a bewildered parent dealing with a bad child. Despite past failure, the only thing the parent knows to do is to admonish him oftener and spank him harder. A wiser procedure would be first to try to understand what is going on in the child's head. Frequently the child does not *want* to misbehave, and most assuredly the driver does not *want* to kill himself.

Intelligent measures toward educating the public to drive safely can only

[9] Note that our definition of steering obviates the necessity of any 'rule' for getting out of a skid.

be taken when the performance of driving an automobile is thoroughly understood. Programs of testing and experimenting are necessary, but useful only if they have some theory to go on. Discussion of the problem should be disinterested rather than merely admonitory, should adopt the driver's point of view rather than the safety engineer's, should start from normal rather than abnormal driving, and should emphasize what the driver ought to do rather than what he ought not to do.

If driving is to be taught successfully—and there are already courses in a number of adult education and public school systems—it must be taught in terms which the driver will recognize. The concepts employed must refer to operations which the driver understands. If teaching is to be more than merely the memorizing of the local legal restrictions, the 'causes' of accidents as listed in statistical tables, a list of 'bad driving' practices, and a series of assorted 'don'ts,' then a systematic theory of driving is needed which can be verified by observation and experimentation. The theory which has been described is an effort in this direction.

Safe and efficient driving is a matter of living up to the psychological laws of locomotion in a spatial field. The driver's field of safe travel and his minimum stopping zone must accord with the objective possibilities, and a ratio greater than unity must be maintained between them. This is the basic principle. High speed, slippery roads, night driving, sharp curves, heavy traffic, and the like are 'dangerous,' when they are, because they lower the field-zone ratio. Hidden obstacles are dangerous, when they are, because they tend to put the driver's field of safe travel out of correspondence with reality. The various 'bad driving practices' are bad, when they are, for a *reason* and not simply because the traffic cop says so. The reason lies in the psychological principles; the 'rules of the road' follow from these principles and it is the principles, not the rules, which need to be taught.

The skillful driver recognizes the valences of obstacles quickly and automatically and projects their clearance-lines correctly. He knows the boundary line of a potential skid on the curve ahead in precise relation to his speed, the road surface, and the characteristics of his car. A sudden shift of the events in his field will lead to an immediate reorganization of the field-relations and not to a disintegration of them in panic. He calculates the other driver's field of travel with sagacity and never allows it to overlap his own. Finally, he does all these things because he has learned to do them, not because he is frightened into a continual state of strained attention.

# A QUALITATIVE STUDY OF THE EYE MOVEMENTS OF GOOD READERS

By ROBERT Y. WALKER, Ohio State University

An earlier paper by the author dealt with reading time, number of fixations per line, pause-duration, extent of forward shift, and regressions in the eye movements of good readers.[1] These variables of the mechanics of reading are undoubtedly indices of other subjective factors of reading and are probably the result either of early methods of training, such as oral reading practice and 'word' reading, or of comprehension of the reading material. Comprehension of written ideas depends among other factors upon the breadth of vocabulary, extent of knowledge of the reader, variety of past experiences, and understanding of the finer shades of meaning as determined by the syntax of the sentence.

Many investigators of reading ability believe that the rapid reader is able to grasp the context of a line of printed material with two or three fixations rather evenly spaced in the line of reading. In an attempt to support or refute these contentions the records obtained from the earlier investigation were reëxamined and an analysis was made of the distribution of time, number of fixations, and length of fixations over the entire line of reading material as shown by good readers. Factors that might show the cause of the extent of the shift, the duration of the fixation, or the possible cause of regressions were also considered.[2]

*Subjects.* The *S*s were 50 students in the entering class of the University of Iowa. They were selected on the basis of their rating in the top decile of their class in the Iowa Silent Reading Test and the top decile of the Intelligence Test.

*Method and procedure.* Photographic recordings by means of corneal reflection of their eye movements while reading six selections were obtained. Selections I, II, and III were of increasing difficulty. Selection I was a paragraph from a second grade reader; Selection II, a paragraph from an elementary college text on psychology; and Selection III, a paragraph from a very difficult treatise on psychophysiology. In so far as possible equal comprehension was required for the three selections. Four true or false statements were submitted to the *S* on the content of each selection: two statements based on ideas expressed in the reading material, and two statements answers to which could be inferred from the material. Selections II, IV, and V were of equal difficulty as regards vocabulary and reading. Comprehension for Selection IV required merely an idea of the subject matter being discussed. Comprehension

* Accepted for publication July 12, 1937.

[1] R. Y. Walker, The eye movements of good readers, *Psychol. Monog.*, 44, 1933, (no. 199), 95-117.

[2] A review of the literature bearing on this problem was included in the earlier paper.

for Selection II was covered by the four true or false statements as explained above. Comprehension for Selection V required as detailed a report of the content as *S* could give. Selection VI, equal in difficulty to Selections II, IV and V, was used as an examination of skimming. *S* was required to examine the paragraph until he found the material which would answer a question that he had been asked. A comparable selection preceded each recording to familiarize *S* with the requirements.[8]

*Problems.* Two problems were considered: (1) the effect of mechanical difficulty of the paragraph as set up by the construction, vocabulary, and subject matter; and (2) the subjective factors concerning the requirement of comprehension with almost identical mechanical conditions. Little could be determined about a possible third problem on the techniques of skimming due to the short length of the paragraph and the variety of methods used by the *S*s. In using the phrase "mechanical difficulties" of the paragraph, the author realizes that the same paragraph probably presents a different subjective problem to every reader. At present no way of avoiding this variable has been found inasmuch as the hereditary or environmental factors affecting the *S*s could not be controlled.

*Results. (1) Distribution of time and of fixations.* To examine the distribution of time and fixations throughout the entire line it was fractionated into units of 5%.

The number of fixations and the duration of such fixations falling in each 5 percentile were plotted. No rigid quantitative examination could be made since the number of lines in each paragraph of reading material ranged from 6 to 18 lines.

There was no pronounced evidence of the reader spending more time on one part of the line than on another, regardless of the difficulty of the material being read or the requirement of comprehension. Fig. 1 shows the distribution of time for the 6 selections. The solid line shows the regression from a horizontal axis for the time spent in the 5 percentile units of the line. The duration of all fixations occurring in every unit was added for all cases and the mean was determined for the number of *S*s who had fixations in that unit. All values were equated for number of lines in the paragraph and were reduced to a common factor of one line. There was a tendency to spend more time over the entire line in Selection V which required complete comprehension. This was the result of increased frequency and duration of fixation. As was to be expected, skimming required less total time than any of the other paragraphs, the time being distributed equally over the entire line of reading material.

[8] Samples of all reading paragraphs are contained in *Appendix A*, which is included in the thesis on file in the library of the University of Iowa.

Similar results were found for the average number of fixations made in each 5 percentile unit of the line. Fig. 2 shows the distribution of the

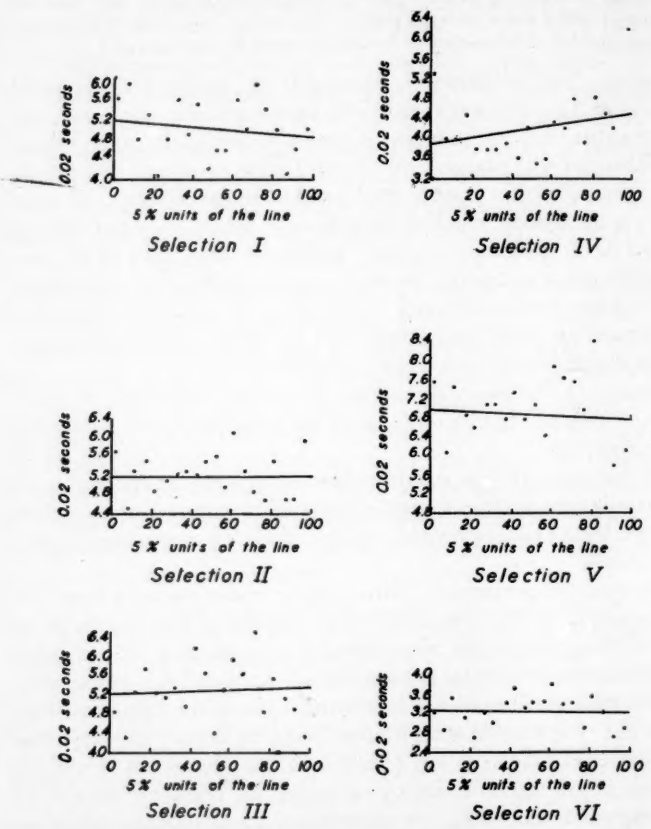


FIG. 1. AVERAGE AMOUNT OF TIME SPENT PER UNIT BY 50 GOOD READERS IN 5 PERCENTILE UNITS OF THE LINE

The points indicate the average time per unit in 0.02 sec. equated to one line of reading. The solid line is the regression curve for the values.

number of fixations per unit over the line of reading material for each of the 6 selections. The solid line is the regression value for the selection. The number of fixations occurring in each unit of the line was deter-

mined and the mean for all *Ss* with fixations in that unit was plotted. The values were corrected for a common factor of one line.

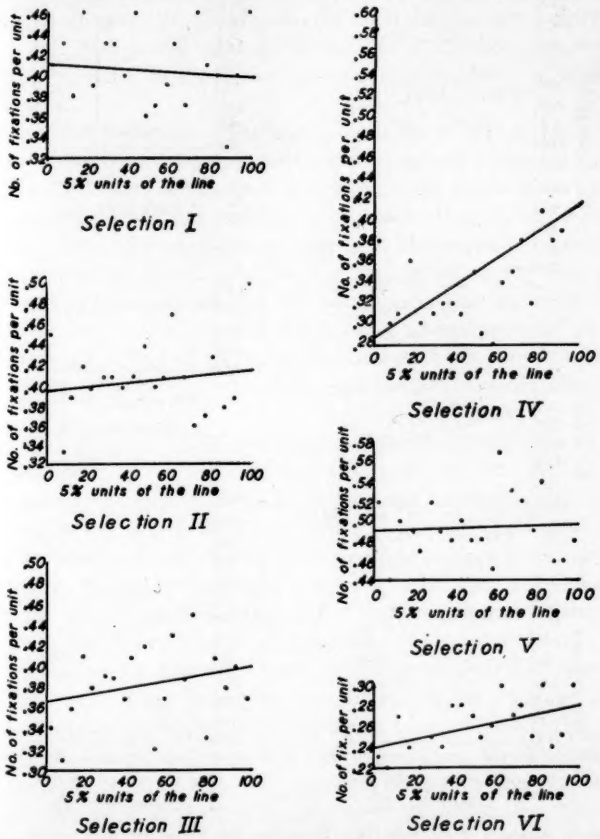


FIG. 2. AVERAGE NUMBER OF FIXATIONS PER UNIT OF 50 GOOD READERS IN 5 PERCENTILE UNITS OF THE LINE

The points indicate the average number of fixations per unit equated to one line of reading material. The solid line is the regression curve for the values.

The average duration for each fixation of the good reader did not remain constant throughout the entire length of the line. Fig. 3 shows the values for average duration per fixation for the 6 selections of reading ma-

terial. There was a tendency for this variable to be slightly longer in the first part and immediately following the center of the line. A noticeable

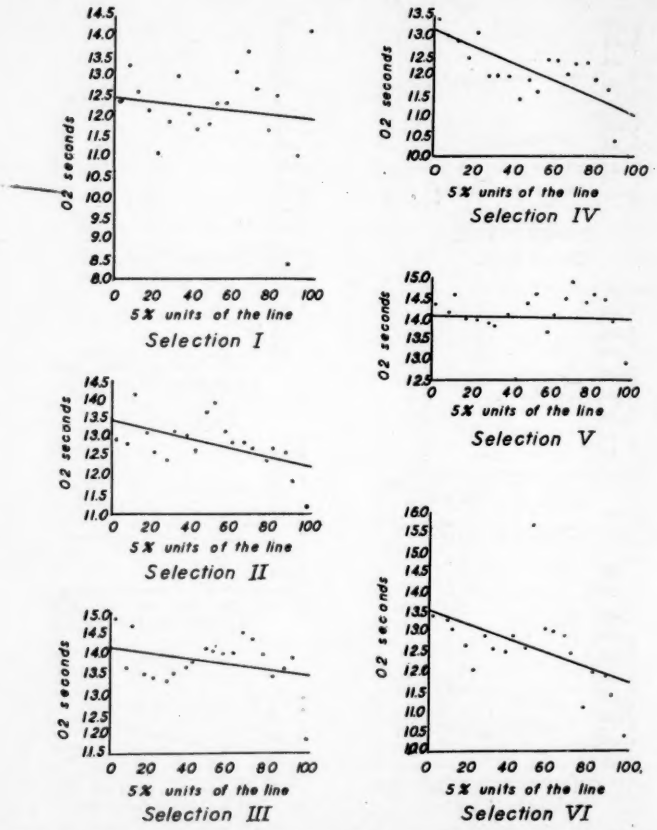


FIG. 3. AVERAGE DURATION PER FIXATION PER UNIT OF 50 GOOD READERS IN 5 PERCENTILE UNITS OF THE LINE

The points indicate the average duration-time in 0.02 sec. per unit. The solid line is the regression curve for the values.

reduction in the average duration per fixation was present at the end of the line in all selections but Selection I, which was very easy reading material with the requirement of moderate comprehension. The increase in

time at the first of the line was probably caused by slight changes in accommodation following the return sweep. A very short interval of time was apparently spent on checking of the line and adjusting the eyes to the new position. No explanation can be given for the slight increase at the middle of the line. The general reduction at the end of the line was probably caused by the small amount of material to be observed when fixations fall at this point.

These results show that differences in difficulty of reading material or changes in the requirement of comprehension do not produce a marked change in the amount of time spent on any part of the line or in the number of fixations. With each line fractionated to such small units there is no evidence of a habit of moving the eyes any fixed distance or in any rhythm. If the good reader moved across the line in relatively equally sized shifts and started at approximately the same point, such small regression coefficients would not be found. If there were habits of eye movement in reading, they must be looked for in the duration of the fixation. There was very slight evidence of such a possibility, but the differences are not significant.

As far as the good reader was concerned, habits did not seem to have been established for point of fixation or for duration of fixation. The ideal reader should be ready and able to vary his eye movements and not be hampered by any system. Such fixed habits of reading would tend to increase the number of regressions. This would be especially true in difficult material. A significant increase in regressions occurred in Selection V, which had very complete requirements of comprehension comparable to those that are found only in extreme situations.

*(2) Spread of the fixations.* Closely related to the preceding discussion is the percentage of each line covered by the fixations, the percentage of indentation, and the percentage of the remaining margin. Table I gives these values for the 6 selections of material used in this study. These figures may not appear to coincide with the statement concerning the number of fixations and the amount of time spent over the entire line. Since the paragraphs varied from 6 to 18 lines per paragraph, it is quite possible that the *S*s made very small indentations in one line with the large margins remaining at the end or with large indentations accompanied by a small margin at the end.

Aside from Selection VI, a record of skimming, the differences in these three variables were not significant. All differences shown were less than one em of printed material. Again, there were no differences of behavior

on the part of the good reader produced by changes in difficulty of the material being read or by the requirement of comprehension. The difference between amount of indentation and the amount of the remaining margin

TABLE I

AVERAGE PERCENTAGE OF LINE COVERED BY FIXATIONS, OF INDENTATION, AND OF REMAINING MARGIN

| Selection | Line | Indentation | Margin |
|---|---|---|---|
| I | 81.2 | 10.3 | 8.4 |
| II | 80.2 | 10.8 | 8.5 |
| III | 79.4 | 10.6 | 9.4 |
| IV | 79.0 | 13.2 | 7.4 |
| V | 79.4 | 10.2 | 10.5 |
| VI | 74.5 | 14.7 | 11.3 |

was not large enough to be significant. This is contrary to the common idea that we see more of the material to the right of each fixation. If such were true the indentation should be smaller than the margin remaining at the end of the line.

*(3) Extent of shifts vs. duration of fixation.* Any relationship that might exist between extent of forward shift and duration of fixation was examined by correlating the duration of fixation with the extent of shift preceding fixation and the extent of shift following fixation. In case the extent of shift had any influence upon the duration of fixation, it should be present in the first correlation; or, if a long fixation permitted clearer examination of adjacent material, we might expect a large shift following such a fixation. Table II gives the coefficients of correlation for the three variables. Correlations obtained were not significant, ranging from —0.08 to +0.30 in the case of extent of preceding shift against duration of following fixation, and from —0.15 to +0.15 regarding the extent of following shift

TABLE II

COEFFICIENT OF CORRELATION BETWEEN DURATION OF FIXATION AND THE EXTENT OF THE PRECEDING AND FOLLOWING SHIFTS

| Shift | Selections | | | | | |
|---|---|---|---|---|---|---|
| | I | II | III | IV | V | VI |
| Preceding | +.30 | +.07 | +.02 | −.08 | +.04 | +.16 |
| Following | +.14 | .00 | +.02 | +.15 | +.04 | −.15 |

against duration of preceding fixation. Different values might have been found if specific fixations had been selected, but a general sampling was desired and no attempt was made to select the fixations used in the correla-

tion on the basis of a difficult or easy word. A consistent sampling of the third fixation in the fourth line was taken throughout the selections. In case a regression occurred at the point selected that fixation was not included in the correlation.

*(4) Mechanical factors.* In attempting to find a mechanical factor that might be the cause of regressions, all forward shifts were segregated into four categories: (A) all forward shifts, (B) forward shifts preceding regressions, (C) forward shifts following regressions, and (D) all others or pure forward shifts. Shifts following regressions were classed separately, for the reader had observed this material and might go beyond the last point of regard when he continued along the line. The means of Category

TABLE III

MEAN AND STANDARD DEVIATION (IN EMS) OF FOUR CATEGORIES OF FORWARD SHIFTS

| Category | | Selection | | | | | |
|---|---|---|---|---|---|---|---|
| | | I | II | III | IV | V | VI |
| (A) All forward shifts | M. | 4.16 | 4.28 | 4.07 | 4.52 | 3.89 | 4.34 |
| | S.D. | .88 | .76 | .53 | .83 | .51 | .80 |
| (B) Shifts preceding regressions | M. | 4.71 | 4.59 | 4.40 | 4.67 | 4.12 | 4.35 |
| | S.D. | 1.34 | 1.01 | 1.01 | 1.66 | .79 | 1.24 |
| (C) Shifts following regressions | M. | 5.88 | 5.69 | 5.32 | 5.54 | 4.68 | 5.27 |
| | S.D. | 1.12 | 1.44 | .95 | 1.49 | .94 | 1.68 |
| (D) All other forward shifts | M. | 3.44 | 3.89 | 3.65 | 4.40 | 3.59 | 4.13 |
| | S.D. | .73 | .75 | .45 | .85 | .47 | .87 |

A, including all forward shifts, ranged in extent from 3.89 ems to 4.52 ems for the six selections of reading material. Table III gives the values in ems for the four categories listed above. The means of forward shifts that preceded regressions, Category B, ranged in size from 4.12 ems to 4.71 ems; the means of shifts following regressions, Category C, ranged from 4.68 ems to 5.88 ems; the means of "pure" forward shifts, Category D, ranged from 3.44 ems to 4.13 ems.

Critical ratios determined between the measures in these categories for the six reading selections are shown in Table IV. In no case was there a significant difference between all forward shifts and those that preceded regressions. If, however, forward shifts, both preceding and following regressions, were removed from the total to leave "pure" forward shifts, a different picture is presented. In skimming or in case a very slight comprehension was required, the critical ratios between "pure" forward shifts and forward shifts preceding regressions was not significant. In all other

situations the critical ratios were highly significant. This was due to the increase in size of the forward shift following the regression. The critical ratios between forward shifts following regressions and any of the other three categories was significant in all but one case. Here the reading material was moderately easy and the requirement of comprehension very slight. In addition very few regressions occurred in this selection which reduced the importance of the measure for this one case.

The above results indicate that an increase in the extent of forward shift is a causative factor in regressions. The large increase in extent following a regression was not unnatural because the subject probably retained

TABLE IV

THE CRITICAL RATIOS AMONG THE FOUR CATEGORIES OF FORWARD SHIFTS FOR SIX SELECTIONS OF READING MATERIAL

| Categories | Selection | | | | | |
|---|---|---|---|---|---|---|
| | I | II | III | IV | V | VI |
| A—B | 2.31 | 1.74 | 2.07 | .53 | 1.70 | .02 |
| A—C | 7.96 | 6.09 | 8.10 | 4.03 | 5.11 | 3.50 |
| A—D | 1.34 | 2.52 | 4.17 | .74 | 3.10 | 1.26 |
| B—C | 4.37 | 4.28 | 4.64 | 2.65 | 3.16 | 3.08 |
| B—D | 3.39 | 3.86 | 4.76 | .99 | 4.03 | 1.00 |
| C—D | 9.49 | 7.68 | 11.27 | 4.50 | 7.21 | 4.22 |

memory traces of the material observed on a previous examination of this point. In such a case the rapid reader would go beyond this point.

Some of the results of this study point toward the possibility of regular habits of eye movements in reading. If such be the case, the good reader has habits quite adaptable to varying types of material and requirements of comprehension. Some papers have referred to "short-lived motor habits" in reading. If the eye movements involved in reading were habits, it seems that the reader would be much less adaptable to such changing conditions. Anderson[4] has found that the poor reader is not so adaptable as the good reader used in this study. Most of the deviations from normal observed in the good reader occurred under changes in the requirement of comprehension. This points to comprehension as a very significant determiner of eye movements. In conclusion eye movements of good readers seem to be symptoms of the ease or difficulty in the ability of the reader to comprehend the material.

[4] I. H. Anderson, Studies in the eye movements of good and poor readers, *Psychol. Monog.*, 48, 1937, (no. 215), 1-35.

## Conclusions

*(1)* The good reader distributed his time almost equally over the entire line of reading material.

*(2)* The good reader distributed his fixations evenly over the line as a whole.

*(3)* Differences in difficulty of material or in the requirement of comprehension did not change these results significantly.

*(4)* The duration of fixation was greater than average at the first and the middle of the line.

*(5)* In most cases the duration of fixation was reduced at the end of the line.

*(6)* Amount of line fixated, indentation, and remaining margin were constant under all conditions.

*(7)* Extent of forward shift did not influence duration of fixations, nor did duration of fixation influence extent of shift.

*(8)* Increase in extent of forward shift was apparently one of the causes of regressions.

*(9)* It appears that the eye movements of good readers are symptoms of the ease or difficulty in the ability of the reader to comprehend the material.

# A COMPARISON OF COLOR-BLIND TESTS

By B. R. PHILIP, La Salle College

It is now known that there is not an abrupt dichotomy dividing the color-blind from those with sound vision, but that something like a normal distribution, skewed at the weak end, exists in the perception of color.[1] Hence some method of determining the degree of chromatic sensitivity that will detect color defects with certainty and speed is of value. The exploration of color perceptive ability by means of apparatus is slow and has not the general effectiveness of some of the commoner non-apparatus tests. This paper presents a study of the agreement of several such tests and the extent to which they indicate types of color defectives.

## METHOD AND PROCEDURE

*Subjects and tests.* From among some 500 men college students who had taken the author's color perception test those who gave evidence of color weakness were selected and given the Ishihara test. Those *Ss*, 42 in number, who made one or more mistakes on the Ishihara were then given the remaining tests of the battery which is as follows.

*(1) Ishihara.* Old (sixth) edition, a pseudo-isochromatic test of 16 plates.

*(2) Edridge-Green.* Contains 24 plates resembling the Ishihara in design with letters replacing the numbers.

*(3) Nagel.* Contains 20 plates on each of which is printed a circle of dots. In Part A, consisting of 16 plates, those are to be selected which contain (a) red or reddish dots, (b) only red dots, (c) only green dots, (d) only grey dots. In Part B, 4 plates, the colors on the plates are to be named and pointed out.

*(4) Holmgren wools.* From 125 skeins of yarn of different colors those skeins which match red, pink and green, respectively, are to be selected.

*(5) Philip color perception test.* Forty cards, each containing 36 dots with varying numbers colored blue, green, purple or red, are to be sorted according to the predominating color.[2]

*(6) Ishihara, new edition, 32 plates.* Scores were obtained from 24 *Ss* only, as the test was available only towards the end of the investigation.

These tests were all given by the author in good daylight, under approximately the same conditions; the directions were faithfully followed.

*Method.* After the tests were given and the judgment on each card or skein was recorded it was felt that a method of transmuting the judgments into numerical scores was required. This was done as follows. For all tests except the Holmgren

* Accepted for publication August 25, 1937.

[1] I. A. Haupt, Tests for color blindness: A survey of the literature with bibliography to 1928, *J. Gen. Psychol.*, 3, 1930, 222-267.

[2] B. R. Philip, The discrimination of color mass, this JOURNAL, 51, 1938, 127-135; Sex differences in the perception of color mass, *ibid.*, 398-404.

every error in calling a card (a divergence from what normals are supposed to do, as given by the directions) was counted as one mistake; all such mistakes were totalled for every test. No difficulty was experienced except for the Edridge-Green, which is so difficult that even normals are prone to make mistakes particularly when there are two letters on a card. An omission of one letter was nevertheless counted as a mistake, and a score equal to or better than that made by a group of 10 normal *Ss* was taken as a normal score.

In the Holmgren wool test so much latitude is given by the directions as to what constitutes a correct matching that it was finally determined to score all selections with leniency, particularly in matching the pink skein. The numerical score eventually used was the percentage of mistakes made to the total number of skeins selected, for the green and red standards only. This seemed to be the only feasible way as the number of mistakes varied according to the closeness of matching that *S* fancied was permitted. In the Philip test the total number of errors in 3 trials was recorded —the usual way of scoring. Owing to the coding on the reverse side of each card scoring was simple and definite.

*Types of color deficiencies.* Most writers divide their *Ss* according to color perceptive ability into color blind, who are in turn sub-divided into deuteranopes and protanopes according as they are predominantly green or red color blind, anomalous trichromates, and normals. Following Von Kries, quoted by Mary Collins,[3] the following definitions are useful.

*Deuteranopes:* Those whose color system is reduced to yellow and blue but who see color throughout the whole length of the spectrum.
*Protanopes:* Those whose color system is reduced to yellow and blue but whose spectrum is shortened at the red end. For them the point of maximum brightness is shifted to the green.
*Anomalous trichromates:* Those who see the fundamental colors of Helmholtz but who are unequally sensitive to red and green.

An extra group, the color weak, was added to include all those whose color deficiency is detected by the Ishihara test.

*Criteria of classification.* As the directions in the tests are often vague and an *S* may be differently classified by the various tests, for practical purposes the *Ss* were classified by the following criteria.[4]

*(a) Color blind*

In the *Ishihara test* both color blind types make 11 or more errors on the 16 plate edition and 27 or more errors on the 32 plate edition. The deuteranopes recognize the first digit only and the protanopes the second digit only on Plates 12 and 13 of the old test, and on Plates 22, 23, 24 and 25 of the new edition. The protanopes often call Plates 2 and 3 of both editions correctly; the deuteranopes always call them wrongly.

On the *Edridge-Green test* both types call half the plates at least erroneously. The protanopes see no letter on Plates 1 and 2 and call Plates 7 and 13 correctly; the reverse holds for the deuteranopes.

---

[3] Mary Collins, *Colour Blindness*, 1925, 12.
[4] Haupt uses a similar pragmatic division for scores on the Nela Test (I. A. Haupt, The Nela test for color blindness applied to school children, *J. Comp. Psychol.*, 6, 1926, 291-303).

On the *Nagel test* there is no clear cut distinction between the types. They both make errors but the protanopes particularly are wrong in Part A, 2, the calling of the red dot, and in Part B, Plates 2 and 4, the differentiation of the red and brown dots.

In the *Holmgren test* the color blind make glaring errors in the selection of the skeins, the protanopes being particularly weak in matching the red skein.

In the *Philip* test the total number of errors made by the color blind in 3 trials is at least 38. The protanopes make a disproportionately large number of mistakes by calling the cards purple or red.

*(b) Anomalous trichromates*

On the *Ishihara* and *Edridge-Green* tests the anomalous trichromates make less mistakes than the color blind and more than 2 errors on both tests. They nearly always do the *Nagel* and the *Holmgren* tests successfully, and make 35 or more errors on the *Philip* test.

*(c) Color weak*

The color weak make 1 or 2 mistakes on the *Ishihara* and the *Edridge-Green* tests, none on the *Nagel* or the *Holmgren,* and exceed 31 mistakes on the *Philip* test. They do not make the color judgments with the speed and certainty of most normals.

## RESULTS

These somewhat arbitrary criteria have been used in Table I, which gives the frequency distribution of the various types of color deficiencies obtained from our battery of tests. All the *Ss* indicated as color blind or

TABLE I

DISTRIBUTION OF 42 *Ss* ACCORDING TO TYPES OF COLOR DEFECT

| Test | Color blind | | Anomalous trichromates | Color weak | Normal |
|---|---|---|---|---|---|
| | deuteranopes | protanopes | | | |
| Ishihara | 18 | 6 | 11 | 7 | — |
| Edridge-Green | 18 | 6 | 5 | 2 | 11 |
| Nagel | 8 | 11 | 6 | 3 | 14 |
| Holmgren | 9 | 9 | 5 | — | 19 |
| Philip | 19 | 9 | 9 | 4 | 1 |

anomalous trichromates by the Nagel and the Holmgren tests are detected as such by the other tests. On the other hand these two tests fail to detect a considerable percentage of the *Ss* with defective color vision. The other tests are in substantial agreement but the Edridge-Green sometimes fails to detect an anomalous trichromate. The Ishihara test never fails to do so.

The ratio of protanopes to deuteranopes found by the Ishihara and the Edridge-Green tests in this investigation, 6 to 18, is but a first approxima-

tion.[5] Naturally the number of *Ss* is too small to form accurate conclusions. Moreover these two tests seem designed to detect the deuteranopes. On the Ishihara test the old edition has but 2 plates of 16 and the new edition, but 4 of 32, that distinguish the two types. Plates 2 and 3, the same in both editions, may give additional clues. Besides the distinction between the two types is not a hard and fast one for the color weak region of the spectrum may vary considerably with the individual *S*, making classification difficult. In general, however, one may state that the deuteranopes greatly outnumber the protanopes.

It is not to be inferred that those deficient in color perception always discriminate colors less effectively than do normals. Their discrimination is different and certain color combinations may even stand out far more sharply for them than for normals. Hence it is that on Plates 10 and 11 of the old edition (Plates 18 and 19 of the new edition) and on Plates 20 and 21 of the new edition, they may see the figures quite clearly. This certainly is not a color defect; rather does it indicate greater perceptive power than normal: yet it is scored as an error because it does indicate an anomaly in color perception.

*Comparison of the Ishihara tests.* The old edition, (16 plates) was given to all *Ss*. The new edition, (32 plates) appeared while the investigation was in progress and was given to 24 *Ss*. Thus a comparison could be made between the two editions. The correlation coefficient is 0.95 which is quite high, and probably is a representative measure of the reliability of the test. On the other hand 15% of the judgments on the same plates were different. This is not surprising as an *S* might at times name different numbers on a plate even within the course of a few minutes. Many of the judgments to be made were just at the threshold of discriminative ability for color weak *Ss* and hence for them subsidiary factors, which have little or no influence on the quick, decisive judgments of those competent in color, assume important rôles. Such factors are the illumination incident on the plate, the angle it is held before the eye, and fatigue and attentive set of *S*. A color weak *S* makes color judgments slowly and doubtfully and his scores do not have the reliability of those competent in color.

The impression was gained throughout that the new edition is harder even for normal *Ss* than is the old. The plates differ, but slightly it is true, in hue and saturation from the old edition and the figures do not emerge with the same clearness. In confirmation of this impression is the fact that the 24 *Ss* made an average of 12% more mistakes in the new edition when similar plates of both editions were judged. Less errors were made on Plates 19 and 26 of the new edition, corresponding to Plates 11 and 14 of the old edition. This is evidently due to the fact that the plates of the old edition are much more legible; as the perception of the figures on these plates is scored as an error it consequently is more often erroneously judged.

[5] Similar ratios have been found in previous investigations; *e.g.* W. R. Miles, 100 cases of color blindness detected with the Ishihara test, *J. Gen. Psychol.*, 2, 1929, 538.

*Intercorrelations of the tests.* One of the objects of changing the results into scores was the securing of Pearson product moment correlations.[6] Two such tables of intercorrelations are presented. Table II gives the results for 35 *Ss* who are color blind or anomalous trichromates. Table III

TABLE II

INTERCORRELATIONS OF COLOR BLIND TESTS
(Data from 35 color blind or anomalous trichromate Ss.)

| Test | Ishihara | Edridge-Green | Nagel | Holmgren |
|---|---|---|---|---|
| Ishihara | — | | | |
| Edridge-Green | 0.79 | — | | |
| Nagel | 0.63 | 0.56 | — | |
| Holmgren | 0.53 | 0.65 | 0.79 | — |
| Philip | 0.48 | 0.60 | 0.53 | 0.48 |

gives the intercorrelations obtained from 42 *Ss*, thus including the scores of 7 color weak *Ss*. As these *Ss* made but one or two errors on the Ishihara test and no errors on every other test but the Philip, the correlation coefficients are of course greatly increased. They are, however, hardly affected for the Philip test because this is a test of both normals and color blind, and the scores of the additional 7 *Ss* are not massed at one end of the distribution as in the case of the others.

TABLE III

INTERCORRELATIONS OF COLOR BLIND TESTS
(Data from 42 Ss with defective color vision.)

| Test | Ishihara | Edridge-Green | Nagel | Holmgren |
|---|---|---|---|---|
| Ishihara | — | | | |
| Edridge-Green | 0.90 | — | | |
| Nagel | 0.70 | 0.87 | — | |
| Holmgren | 0.61 | 0.54 | 0.80 | — |
| Philip | 0.50 | 0.50 | 0.53 | 0.50 |

The correlation coefficients are all positive and range from 0.5 to 0.9. Such values are high in spite of the arbitrary method of scoring the tests, the fewness of *Ss*, and the differences in material and methods of the tests. Certainly color defect is a trait that may be measured consistently though with some degree of variability. Had normal *Ss* been included the intercorrelations and reliability of the tests would have been greatly increased. As the purpose of this investigation, however, was chiefly to determine the value of the tests in differentiating degrees of color defect no normals

[6] S. W. Terman, (A new classification of the red-green color-blind, this JOURNAL, 41, 1929, 237-251) reports coefficients of 0.82 between the Ishihara, the Stilling and her own.

were included. The dispersion of the scores and hence the intercorrelation coefficients are thereby diminished.[7]

*Relative merits of the tests.* The *Ishihara test* is by far the most effective test of color blindness which does not involve apparatus. It is rapid, taking only from 1–5 min; it detects with certainty color blind and anomalous trichromates, and furnishes convincing evidence even to the *S*s of their anomalies. If anything the older edition is more effective as it is passed with ease by normals who may experience difficulty with some of the plates of the new edition. On the other hand the Ishihara test does not always detect color weak *S*s. Several *S*s passed the Ishihara without an error but made a few errors on subsequent re-testing. The Ishihara does not give any clue to the wide variability of color perception in normals nor does it clearly indicate the color confusions made.

The *Edridge-Green test* is also effective in discriminating the color blind and the anomalous trichromates, but it is somewhat slow, taking from 5–15 min. Unlike the Ishihara it has no plate readily passed by all the *S*s (Plate 1 of both editions) to demonstrate the size and general form of the letter or figure to be perceived. It offers considerable difficulty to normals and does not show to color defectives their anomalies in a convincing fashion. The *S*s may vary considerably in being re-tested by it.

The *Nagel test* is slow, taking from 5–20 min., and is quite ineffective, as many trichromates pass it without an error. Hence it is not recommended except for confirmatory use.

The *Holmgren test* is very slow, taking from 5–30 min. It does not detect all anomalous trichromates. Its chief advantage is to show the types of color confusions made.

The *Philip test* is not properly speaking a color-blind test. It does indicate the color perceptive ability of all the *S*s and yields a score from which the decile rank of *S* may be ascertained. It is a useful supplement to the Ishihara as it indicates those whose color weakness would otherwise pass unnoticed. For this purpose the following criteria may be used, assuming that the test is given under proper conditions and that *S* coöperates to the best of his ability.

*(a)* Those adults who score 38 or over may be suspected of color blindness or at least of being anomalous trichromates. This score is made or exceeded by the lowest 7% of a distribution of 485 college men. Corresponding scores for anomalous

[7] T. L. Kelley, *Statistical Method,* 1924, 221.

trichromates and color weak are 35 and 32, which were made or exceeded by 11% and 16% of this group.[8]

*(b)* Those who make a considerable proportion of mistakes on the easier types of cards (Types 18 and 16). Roughly 40% or over are made on these types by the color blind, 30% by the anomalous trichromates, and 20% by the color weak. The normal average made by the group of 485 men is 10%.

*(c)* Those whose confusion colors are predominantly blue or red, or any combination involving these, rather than purple or green. The average percentages made by the normal group are as follows: blue, 9%; green, 33%; purple, 50%; and red, 8%. A rough division of the various types, based on the percentages the blue and the red confusions are of the total, is: protanopes, 50–80%; deuteranopes, 40–60%; anomalous trichromates, 30–40%; color weak, 20–35%; normals, 10–25%.

*(d)* Those who take a long time and show considerable hesitancy, particularly on the second and third trials. The normal average time is 340 sec., with a standard deviation of 92 sec. The complete test takes from 6–15 min. Those *S*s who fail by all four criteria are definitely to be suspected of color blindness. A few *S*s make many errors or take a long time, yet are not color defectives by the Ishihara test, but invariably they lack instruction or motivation.

Where the detection of color blindness is required to avoid dangerous accidents, as on the railroads, the Ishihara test is effective in determining those who are to be subjected to further investigation.

Where information is sought in regard to the confusion colors of a color blind *S* the Holmgren test is of value.

Where keenness of color perception is imperative, as in the selection of color workers, the Philip test is helpful, and may be used to supplement the Ishihara.

---

[8] Diagnostic scores for boys are about 5 points higher. The correlation of the scores of 90 boys of defective color vision, who made two or more mistakes on the new edition of the Ishihara test, and their scores on the Philip test is 0.60 ± 0.04.

# BINOCULAR DEPTH CONTRAST AND THE CONDITIONS OF THE BINOCULAR FIELD

By HEINZ WERNER, Wayne County Training School, Northville, Michigan

In an earlier study,[1] it has been established that 'binocular depth contrast' exists. Whatever the interpretation of this phenomenon may be, it is clear that binocular depth contrast occurs as a common feature in ordinary binocular perception. It is the appearance of depth in a binocular non-disparate form, when this form is surrounded by a disparate configuration. In Fig. 1 the lines $a_1a_2$ and $b_1b_2$ represent such a *disparate stereoscopic*

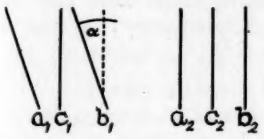


FIG. 1. STEREOSCOPIC LINES WITH DEPTH CONTRAST

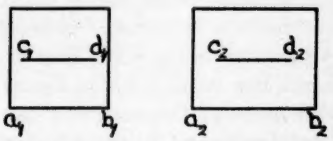


FIG. 2. STEREOSCOPIC FRAME PRODUCING DEPTH CONTRAST IN THE HORIZONTAL LINE

configuration (angle of disparity = $\alpha$). The lines $a_1$ and $a_2$, $b_1$ and $b_2$, binocularly united, produce a depth in the non-disparate (corresponding) center line, $c_1$, $c_2$, which is contrary to the depth in *a* and *b*. That is to say, *a* and *b* are seen rotated in such a way that the top of the line is behind, and the bottom of the line in front of, the plane, while the line *c* stands out with the upper part in front of, and the lower part behind, the plane. The depth effects are qualitatively opposed. The amount of depth contrast varies with the pattern used, with the individual observer, but the phenomenon itself can be demonstrated quantitatively for all *O*s.

In the study previously mentioned an attempt was made to explain binocular depth contrast in a way that would be understood by an inspection of Fig. 2. Most frequently, and especially with untrained *O*s, the binocular frame appears at only a slightly slanting depth, or without any slant at all (frontal-parallel position). This is true despite the considerable disparity of the half-images, according to which one would expect to see the frame turned diagonally, and at a relatively high degree, out of the frontal-parallel plane. If by 'natural' or 'primary' correspondence one means

* Accepted for publication October 25, 1937. From the Harvard Psychological Laboratory.

[1] Dynamics in binocular depth perception, *Psychol. Monogr.*, 49, 1937, (no. 218), 1-120.

that the unification of certain geometric-anatomically defined points in the right and left retinas brings about the impression of single points lying in the fixation-plane, then the laws of this primary correspondence seem to be altered if such a frame-like figure is introduced. In this case the vertical lines, $a_1$ and $a_2$, may be said to correspond with respect to $b_1$ and $b_2$, providing that by the term *correspondence* two characteristics are understood; viz. (1) complete unification of the half-images and (2) appearance of the unified points in the plane of fixation. In other words, should the frame appear frontal-parallel, the correspondence can be considered as having changed from a 'primary' to a 'secondary' one. Moreover, if points of retinal disparity have become correspondent, one should expect, logically, that originally corresponding points belonging to the same field should now function as *disparate* points ('secondary disparity'). If, therefore, anatomically corresponding lines of equal length are introduced into the frames, they should acquire a disparity of a secondary order. This is exactly what happens. The horizontal line *c–d*, instead of lying in the plane, appears rotated on its center with the right end forward. It is a further logical consequence of such a change in correspondence that the rotation of the horizontal line has to be contrary to the inclination that the frame should exhibit according to its retinal disparity.

If this argument is correct, binocular depth contrast might be conceived as a transformation of correspondence occasioned by the introduction of retinally disparate configurations into a binocular field.

There are certain consequences in such an interpretation which need confirmation by means of further experimental evidence. If contrast appears as an effect caused by the shift in correspondence within a binocular field, this contrast effect should be dependent on the apparent depth of the retinally disparate figure (the frame, in our case). That is to say, if there is an increase in apparent depth in the frame (approaching natural correspondence), there should be a decrease in the contrast effect, and vice versa. Now, the reports of the *Os* participating in the original experiment have, in general, been in consistent agreement with that very expectation.

Since confirmation of the underlying theory cannot be left entirely to introspection, in this study it will be shown how the problem has been approached by means of a new series of experiments.

*Apparatus and procedure.* The measurement of the amount of depth contrast was carried out in the same way as in the earlier experiment, *i.e.* by means of the contrast-stereoscope described in the earlier study.[2] The principle of measurement

[2] Werner, *op. cit.*, 106-109.

may be briefly restated. The lines $a_1$, $b_1$, $c_1$ *and* $a_2$, $b_2$, $c_2$ of Fig. 1 are represented in the apparatus by 6 thin black wires. The binocularly unified middle line appears, by contrast, turned out of the frontal-parallel plane. The *O* can counterbalance the contrast effect by turning the right-hand middle line $c_2$ about its center in a clockwise direction. The amount of rotation necessary to bring the unified middle line back into the frontal-parallel plane indicates the amount of contrast. See Fig. 3.

FIG. 3. NECESSARY ROTATION OF $c_2$ ($\angle\beta$)

Since we are interested here in the relation between the contrast and apparent depth in the outer lines *a* and *b*, some sort of objective indication of the angle at which these lines appear seems indispensable. The indicator used in the experiment consists, essentially, of a black needle which is rotated by the *O* around horizontal axis in front of a circular white field. The deviation of the needle from a vertical (zero) position to one side or the other represents the angle at which the *O* sees *a* and *b* standing out of the frontal plane. This angle can be read off a scale on the back of the indicator.[3]

The procedure in each single trial is as follows: First *O* adjusts the right hand middle line $c_2$ in the contrast stereoscope so that the binocularly united line *c* appears in the frontal-parallel plane (angle of sidewise rotation of $c^2 = \angle\beta$). Next, he turns the needle of the indicator until the deviation from the vertical line seems to equal the angle at which the outer lines *a* and *b* appear to have moved out of the frontal-parallel plane ($\angle\gamma$).

*Results.* Since it is the relation between the apparent depth of the outer lines and the contrast depth of the middle line that has to be tested, the easiest way, perhaps, to examine this relation would be to keep the disparity angle $\alpha$ constant, and to plot the variations in the amount of apparent contrast against the variations in the apparent depth of the outer lines, *a, b*. Unfortunately the interdependence is not sufficiently direct and simple to permit such a method of exposition. If the single trials performed at a constant value of $\alpha$ are examined, it may very well be discovered that, as a general rule, an increase in the apparent depth of *a* and *b* is accompanied by a decrease in $\angle\beta$, and vice versa. It is also not an unusual occurrence that, in comparing the single trials with each other, the increase or decrease in the apparent depth of *a* and *b* is hinged to an increase or de-

[3] The indicator was devised by J. Volkmann for the study of problems concerning the relation between depth perception and retinal disparity. This study, performed in the Harvard Psychological Laboratory by J. Volkmann and myself, will be published later.

crease in the contrast depth of *c*. In other words, the outcome, should only the single trials be considered, might very well be ambiguous, that is to say, might be occasioned by not one, but several, factors which are counterbalancing their specific effects.

As for the cause of the parallelism just mentioned between apparent depth in *a, b* and contrast depth in *c*, I am led to believe that the underlying factor can be found in the relative stability, or lability, of fixation. Without inquiring too deeply into this matter, there is a good deal of evidence to support the contention that eye-movements cause an increase not only of the apparent depth in the outer lines, but also, generally speaking, of contrast depth. We already know from certain experiments that the utilization of a given disparity for the perception of depth is facilitated by eye-movements.[4] Now, if it is assumed that contrast depth in the middle line *c* is based upon a 'secondary disparity' (as a consequence of the shift in correspondence), the law of increase in depth through eye-movements may be understood as applying also to contrast depth.

However this may be, in order to solve the problem in question, one has to seek a method different from the one which simply correlates the single variations in apparent depth of a constant *a* and *b* with the variation in contrast effect.[5] This alternative method is based on the generally accepted assumption that, within certain limits, and if all other conditions are kept constant, there is a steady increase in apparent depth concomitant with an increase in retinal disparity. If the apparent depth values ($\angle\gamma$) are plotted against the steadily increasing disparity values ($\angle\alpha$), a somewhat uniform curve may be anticipated, provided that the number of trials for each disparity is sufficient to rule out the effect of irregularity in fixation.[6] Even so, however, there may still be a change in the conditions of correspondence in the binocular field from one point to the other. If such a change occurs, it should come to light in a certain irregularity in the curve, making it relatively more or less steep at successive stages of disparity. These irregularities may, therefore, be understood as indicating changes in correspondence. On the other hand, the increase in contrast depth might be plotted

[4] This, of course, is only a descriptive statement, the theory of which is presented partly in the above mentioned monograph, *op. cit.*, Part I, and partly in a study, also mentioned previously, by J. Volkmann and myself, performed in the Harvard Laboratory.

[5] There remains, of course, the possibility of using the direct method by employing a technique that precludes eye-movements. This implies strict control of fixation which, of necessity, demands a rather complicated arrangement that we did not try to develop at the present time.

[6] The factor of variation in fixation has been very much minimized by the introduction of a fixation point at the center of the middle line $c_1$, $c_2$.

against the disparity angles $\alpha$ of the outer lines in order to compare both curves as to a possible significant relationship in shape and concurrent turns. This last method has been the procedure used for working up the results. The type of experiment is peculiar in that it reveals large individual differences in the perception of the different *O*s, and therefore the results have been treated separately for each of the four *O*s who took part in the experiment.

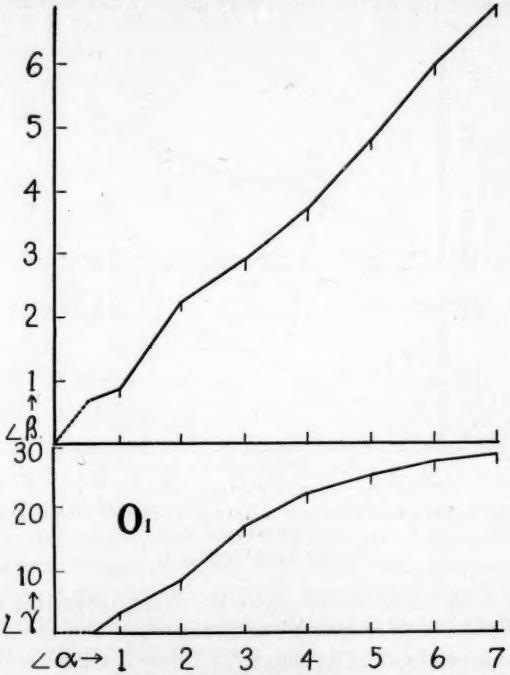


FIG. 4. RELATION BETWEEN APPARENT DEPTH OF OUTER LINES AND CONTRAST FOR $O_1$

The lower graph shows the relation between the disparity angle, $\alpha$, of *a* and *b* and the apparent inclination, $\gamma$. The upper graph shows the relation between the same disparity angle, $\alpha$, and the contrast effect $\angle\gamma$.

The first graph, Fig. 4, is that of *E.G.*, who shows the highest absolute amount in contrast. Disparity ($\angle\alpha$) of the outer lines increases in steps from $\alpha = 0°$ to $\alpha = 7°$. Accordingly there is, in general, an increase in the apparent inclination at which the lines *a* and *b* appear (Fig. 4, lower graph).

At the same time there is also an increase of inclination by contrast in the middle line $c$ (Fig. 4, upper graph). The values for each point are averages taken from 10 trials.

It can be readily seen that the general character of the lower curve shows a somewhat irregular but considerable rise up to about 4° disparity, from which point on there is a very noticeable diminution in steepness. The contrast curve, on the other hand, shows not only a continuation of the preceding slope beyond the point of 4° disparity, but also an increase in

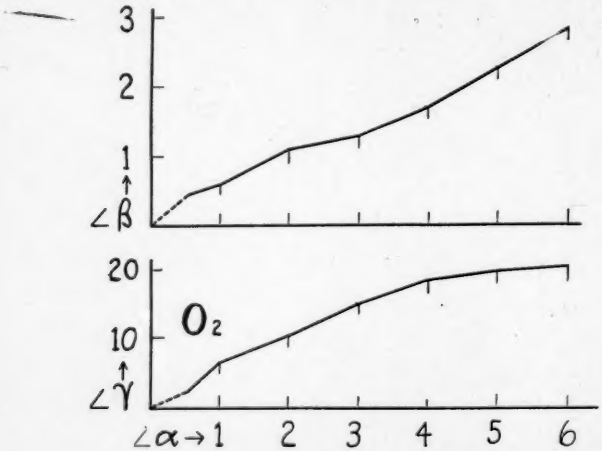


FIG. 5. RELATION BETWEEN APPARENT DEPTH OF OUTER LINES AND CONTRAST FOR $O_2$
(See legend to Fig. 4)

steepness. With respect to this and other turning points the two curves behave almost altogether contrariwise. Where there is an increase in steepness in the lower graph, as at points 30′, 2°, there is a decrease in the upper graph. Where the lower curve becomes relatively less steep, as at points 1°, 3°, 4°, the upper curve becomes steeper. The only negligible exception to this rule is that at the point where $\alpha = 7°$, a rather extreme and therefore, psychologically, not very clearly determined value.

In order to inquire into the validity of the results obtained, the experiment was repeated with the same $O$. Although the values for the apparent inclination, and for the contrast depth as well, are somewhat higher, the general trend observed in the first experiment is still noticeable.

Figs. 5, 6, 7 show the results gained from the experiments carried on with the other three *O*s. A closer inspection of these graphs proves that, notwithstanding the great individual differences, in the apparent perception of depth in *a*, *b* and the amount of contrast effect in *c*, the characteristic relation between each of the two curves is practically the same for all four *O*s.

The results may be summarized as follows. An examination of each pair of curves reveals the following characteristics. (1) With a small disparity (up to 15′–30′), although the outer lines do not appear at any considerable depth inclination, there is nevertheless a relatively steep increase in

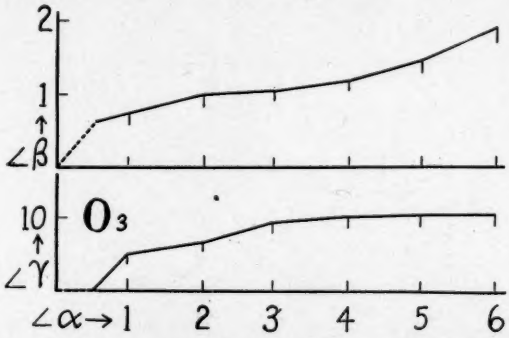


FIG. 6. RELATION BETWEEN APPARENT DEPTH OF OUTER LINES AND CONTRAST FOR $O_3$
(See legend to Fig. 4)

the contrast depth of the inner lines.[7] (2) In higher regions of disparity where the apparent depth curve rises only a little, there is still a very considerable slope of the contrast curve. (3) In general, the direction of the turns of both curves at different values of disparity is more or less noticeably reversed, one with respect to the other.

*Interpretation.* However one may account for the contrast effect, an interpretation which attempts to explain depth contrast as a sort of 'induction,' *i.e.* as being caused directly by the visible depth of the outer lines, would be quite untenable. The mere fact that contrast occurs definitely when the outer lines are perceived without inclination should exclude such an

[7] This statement applies only to 3 out of the 4 *O*s. As to *J.B.*, refer to discussion further along.

argument. As already pointed out, however, certain conditions may cause an increase or decrease in the apparent inclination of the outer lines concurrent with the contrast depth of the middle line. Stability and lability in fixation might be such factors.

Change of correspondence in the binocular field seems to present itself as the basis of an interpretation which, with satisfactory simplicity, can explain the opposition in direction of the turns of the two curves. That

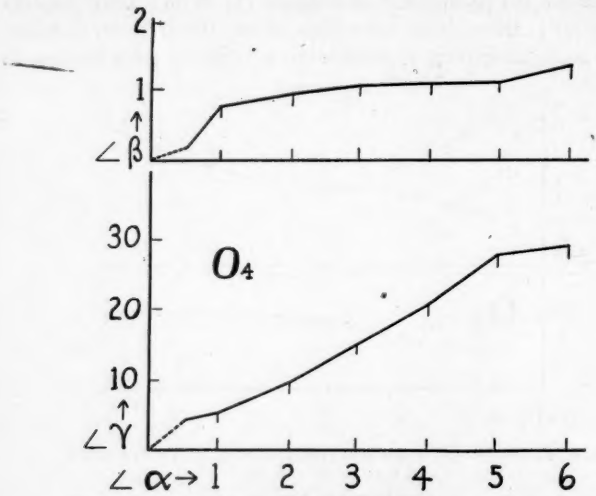


FIG. 7. RELATION BETWEEN APPARENT DEPTH OF OUTER LINES AND CONTRAST FOR $O_4$

(See legend to Fig. 4)

changes in the natural correspondence of the binocular field occur as a consequence of the introduction of certain configuration seems to be a highly probable hypothesis. If we take the term 'change of correspondence' at its purely descriptive value, we can find strong evidence for such a conclusion even if we consider only the beginning of the lower graphs. In this region apparent depth is zero (or almost zero) for three out of four *O*s, although there is present a not inconsiderable amount of retinal disparity.[8] If we assume that this is an indication of a shift away from natural correspondence, a relatively steep increase in contrast effect should be

[8] It could be easily demonstrated that the disparity in question is quite beyond the "physiological" threshold of depth perception.

the logical consequence. Such an increase is actually true of the three *O*s in question.

It is, furthermore, particularly interesting to analyze the graph of the fourth *O* (Fig. 7). His lower curve, in marked contrast to the curves of the three other *O*s, starts with a rather abrupt increase in apparent depth. In a further discrepancy but still in line with the theoretical assumption, his contrast curve shows rather low values within the region of 30′. On the other hand, while the lower curve continues almost horizontally after this value, the upper curve takes an abrupt turn upward.

In other words, when there is a big shift away from natural correspondence (indicated by a relatively negligible or small amount of apparent inclination in the outer lines), the contrast curve rises noticeably, and *vice versa*. The opposition between the two curves at other points can be understood in essentially the same way. We may, therefore, conclude that the findings of this experiment are in keeping with the assumption that stereoscopic depth contrast is the outcome of a change of correspondence within the binocular field.

## Summary

In an earlier experiment the fact has been established that there is a binocular depth contrast, and an attempt has been made to formulate a theoretical interpretation of this phenomenon. The present study tries to test out certain aspects of such a theory by contributing some experimental evidence. An inverse relation has been found between the amount of apparent depth in the configuration which causes the contrast and the amount of contrast depth. These results coincide with the theoretical assumption that binocular depth contrast is caused by a change of correspondence within the particular binocular field.

# CONDITIONED DISCRIMINATION: THE EFFECT OF KNOWLEDGE OF STIMULUS-RELATIONSHIPS

By ERNEST R. HILGARD, RONALD K. CAMPBELL, Stanford University, and WM. NORMAN SEARS, Yale University

Previous studies of conditioned discriminatory eyelid reactions, reported as Experiments I and II, demonstrated that the ability to describe the stimulus-sequences in words did not guarantee conditioned discrimination.[1] On the other hand, there was some indication that the discrimination was more successful for those *Ss* who were able to state early in the course of conditioning that one visual stimulus was always reënforced, the other never reënforced. This verbal knowledge could be related to the motor responses only approximately, however, because the free reports did not reveal the precise point at which the correct hypothesis regarding the stimulus-relationships was accepted.

## EXPERIMENT III. EFFECT OF KNOWLEDGE OF STIMULUS-RELATIONSHIPS

In Experiment III reported here, the *Ss* of one group were told which stimulus was to be negative (not reënforced) and which positive (regularly reënforced). They were thus given full information about the stimuli, except for the order of presentation. The *Ss* of a control group were not given this information. By comparing the course of conditioned discrimination within the two groups, the effect of verbal knowledge of stimulus-relationships can be determined more precisely than in the earlier experiments.

*Procedure.* In order to establish approximately equated experimental and control groups, a first period of experimentation (Day 1) was devoted to the development of simple conditioned eyelid responses to the visual stimulus which was later to be the positive stimulus within the discrimination experiment. This conditioned stimulus was an illumination increase on the left of two adjacent windows, as in the earlier experiments. An air-puff to the cornea of the left eye regularly followed the visual stimulus at an interval of 600 ms. On Day 1 there were presented 60 paired light and air-puff stimuli in groups of 12, spaced at approximately 30 sec., with 2-min. rest-periods between sets of 12. For the following period (Day 2) *S* was assigned

* Accepted for publication November 3, 1937. This investigation was supported in part by grants to the senior author from the Social Science Research Fund of Stanford University, and from the National Research Council.

[1] E. R. Hilgard, A. A. Campbell, and W. N. Sears, Conditioned discrimination: Development of discrimination with and without verbal report, this JOURNAL, 49, 1937, 564-580.

either to the control group (Group A) or to the experimental group (Group B). Assignment to the groups was haphazard; 14 *Ss* served in Group A and 11 in Group B. Records were treated from the 10 *Ss* of each group whose performances during the simple conditioning of Day 1 were most nearly comparable.

If an *S* was assigned to Group A, he received no new instructions at the beginning of the second period. The experiment continued on Day 2 as it had on Day 1, except that on half the trials the illumination change occurred on the right window, which had never been illuminated on Day 1. The light on the right window was never reënforced by the air-puff; light on the left was reënforced as usual. Illumination of the left window was therefore the positive conditioned stimulus during the discrimination series, illumination of the right, the negative stimulus. The negative stimulus was introduced without warning to the *Ss* of Group A.

If an *S* was assigned to Group B, he was informed at the beginning of Day 2 that the right window occasionally would be illuminated, but that it would not be followed by the puff of air, and that illumination of the left window would continue to be reënforced as on the first day. This information given in advance to Group B could become available to Group A only in the course of experimentation.

*Results: (1) Frequency of conditioned responses.* It is possible to determine from the photographic record of each trial whether or not a conditioned response occurred. For this purpose, a conditioned response of 1 mm. or over on the record was counted, and the frequency expressed as the percentage of records on which such responses were found. The gross results for the individual *Ss* of the experiment are presented in Table I.

The *Ss* of each group are arranged in the order of the amount of conditioning shown on Day 1. For this day, the percentage frequency for each *S* is determined from 60 records. The percentage frequency of response to the positive and negative stimuli on Days 2 and 3 is determined from 30 records each. The extent of discrimination is expressed as a difference between the percentage frequency of conditioned responses to the positive stimulus and to the negative stimulus. Where the sign is plus, responses to the positive stimulus are more frequent than responses to the negative stimulus. The final extinction trials are not included in the table, but will be considered later.

The following statements serve to interpret the data reported in Table I. (1) The matching of the two groups on Day 1 is not entirely satisfactory. Group B shows a somewhat higher mean percentage of conditioning than Group A, but the range of values is roughly comparable. (2) Group B shows a higher amount of discrimination than Group A on both Day 2 and Day 3, the proportionate excess being greater on Day 2, when the advantage of information would presumably be greater. The differences are without statistical significance, however, and further analysis is necessary before it may be stated with confidence that the information with regard to the stimuli aided Group B.

*(2) Analysis of responses within early trials.* During the early trials of discrimination the *Ss* of Group A have not had the opportunity of acquiring through experience with the positive and negative stimuli the information given in advance to those of Group B. Differences to be attributed to instructions are probably at a maximum within these early

TABLE I

PERCENTAGE FREQUENCY OF CONDITIONED RESPONSES TO POSITIVE AND NEGATIVE STIMULI THROUGHOUT SIMPLE AND DISCRIMINATORY CONDITIONING
(The *Ss* are arranged in order of frequency of simple conditioning on Day 1.)

| Group | *S* | Simple conditioning Day 1 | Discriminatory conditioning | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | Day 2 | | | Day 3 | | |
| | | Pos. stim. | Pos. stim. | Neg. stim. | Diff. | Pos. stim. | Neg. stim. | Diff. |
| | 44 | 27 | 70 | 27 | +43 | 70 | 17 | +53 |
| | 47 | 33 | 30 | 3 | +27 | 67 | 37 | +30 |
| | 35 | 36 | 76 | 43 | +33 | 70 | 17 | +53 |
| | 29 | 38 | 47 | 20 | +27 | 67 | 27 | +40 |
| A | 50 | 41 | 82 | 64 | +18 | 93 | 68 | +25 |
| (without knowledge) | 40 | 50 | 60 | 33 | +27 | 63 | 10 | +53 |
| | 52 | 51 | 83 | 23 | +60 | 94 | 0 | +94 |
| | 34 | 56 | 79 | 48 | +31 | 93 | 23 | +70 |
| | 56 | 67 | 100 | 97 | + 3 | 97 | 100 | − 3 |
| | 32 | 75 | 85 | 68 | +17 | 77 | 20 | +57 |
| Mean | | 47.4 | 71.2 | 42.6 | +28.6 | 79.1 | 31.9 | +47.2 |
| | 49 | 32 | 57 | 7 | +50 | 47 | 0 | +47 |
| | 37 | 33 | 40 | 7 | +33 | 59 | 21 | +38 |
| | 45 | 38 | 40 | 35 | + 5 | 30 | 20 | +10 |
| | 43 | 40 | 62 | 7 | +55 | 90 | 24 | +66 |
| B | 41 | 66 | 90 | 71 | +19 | 93 | 69 | +24 |
| (with knowledge) | 48 | 67 | 93 | 32 | +61 | 93 | 11 | +82 |
| | 39 | 68 | 57 | 10 | +47 | 93 | 0 | +93 |
| | 42 | 70 | 93 | 73 | +20 | 90 | 72 | +18 |
| | 51 | 77 | 90 | 16 | +74 | 87 | 3 | +84 |
| | 46 | 81 | 86 | 40 | +46 | 90 | 3 | +87 |
| Mean | | 57.2 | 70.8 | 29.8 | +41.0 | 77.2 | 21.3 | +54.9 |

trials. Fig. 1 presents the percentage frequency of conditioned responses to the positive and negative stimuli within each series of 12 trials throughout the three days of experimentation, with the exception of the final extinction trials. The difference between the performances of Group A and Group B within the first series of Day 2 is striking. It is evident that the knowledge given to Group B by instructions has facilitated the discrimination. Within a few trials, however, the groups have become more nearly equivalent. The information given by the experimenter to Group B can be acquired by Group A through exposure to the stimuli, but a number of repetitions are obviously required before the knowledge of the stimulus-

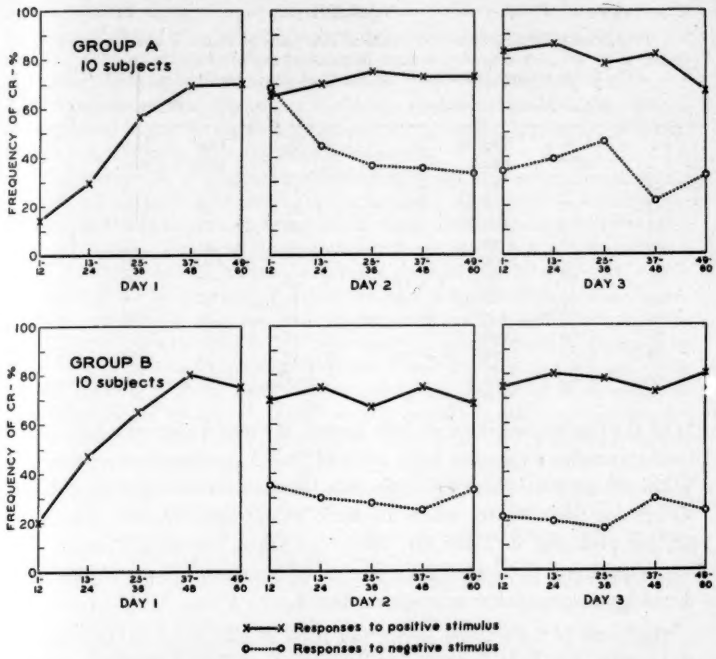


FIG. 1. EFFECT OF ADVANCE KNOWLEDGE OF STIMULUS-RELATIONSHIPS ON THE COURSE OF DISCRIMINATION

Each point represents the mean percentage frequency of conditioned responses from 10 *Ss* within the trials indicated on the base-line. On Day 1 all of the trials were positive (reënforced); on Days 2 and 3 half were positive, half negative (not reënforced). The negative stimulus was introduced within Day 2 without advance warning to Group A, and there was initial generalization. Information regarding the stimuli was given to Group B, however, before the trials of Day 2 began, resulting in immediate motor discrimination.

relationships can be satisfactory. If we assume that knowledge of the positive and negative stimuli is necessary for the responses of Group A to approach those of Group B, it appears that the relationships have been apprehended within the first 24 trials of Day 2.[2]

It is important to examine the data from the early portion of Day 2 more closely, in order to determine the extent to which the individual

[2] For the probabilities of the situation to be reacted to, it may not be necessary for the apprehension of relationships to be at the verbal level. The alternatives will be discussed in connection with later experiments in this series.

TABLE II

PERCENTAGE FREQUENCY OF CONDITIONED RESPONSES TO POSITIVE AND NEGATIVE STIMULI WITHIN TRIALS 1–12 OF DAY 2
(The Ss are arranged in order of frequency of simple conditioning on Day 1.)

| Group A (without knowledge) | | | | Group B (with knowledge) | | | |
|---|---|---|---|---|---|---|---|
| S | Pos. stim. | Neg. stim. | Diff. | S | Pos. stim. | Neg. stim. | Diff. |
| 44 | 67 | 50 | +17 | 49 | 83 | 0 | +83 |
| 47 | 17 | 0 | +17 | 37 | 33 | 0 | +33 |
| 35 | 67 | 100 | −33 | 45 | 33 | 0 | +33 |
| 29 | 17 | 17 | 0 | 43 | 33 | 17 | +16 |
| 50 | 100 | 100 | 0 | 41 | 100 | 83 | +17 |
| 40 | 67 | 50 | +17 | 48 | 100 | 83 | +17 |
| 52 | 67 | 67 | 0 | 39 | 50 | 17 | +33 |
| 34 | 67 | 100 | −33 | 42 | 83 | 67 | +16 |
| 56 | 100 | 100 | 0 | 51 | 100 | 50 | +50 |
| 32 | 80 | 100 | −20 | 46 | 83 | 33 | +50 |
| Mean | 64.9 | 68.4 | − 3.5 | Mean | 69.8 | 35.0 | +34.8 |

*Ss* of the two groups show in their records the tendencies exhibited by the average results. Values for trials 1–12 of Day 2 are presented in Table II. Within these trials the mean difference between the two groups is large. Values for trials 25–36, when the mean values have become more similar, are presented in Table III. These two tables present the crucial data for determining the significance of the differences in discrimination produced by advance information to Group B.

The values in Table II are for the first series of trials in which discrimination was possible. The *Ss* of Group A, not informed about the stimuli, show practically complete generalization. That is, they yield as many conditioned responses to the

TABLE III

PERCENTAGE FREQUENCY OF CONDITIONED RESPONSES TO POSITIVE AND NEGATIVE STIMULI WITHIN TRIALS 25–36 OF DAY 2
(The Ss are arranged in order of frequency of simple conditioning on Day 1.)

| Group A (without knowledge) | | | | Group B (with knowledge) | | | |
|---|---|---|---|---|---|---|---|
| S | Pos. stim. | Neg. stim. | Diff. | S | Pos. stim. | Neg. stim. | Diff. |
| 44 | 100 | 33 | + 67 | 49 | 33 | 0 | + 33 |
| 47 | 33 | 17 | + 16 | 37 | 17 | 17 | 0 |
| 35 | 100 | 33 | + 67 | 45 | 50 | 20 | + 30 |
| 29 | 33 | 0 | + 33 | 43 | 67 | 17 | + 50 |
| 50 | 40 | 50 | − 10 | 41 | 83 | 83 | 0 |
| 40 | 83 | 50 | + 33 | 48 | 100 | 25 | + 75 |
| 52 | 100 | 0 | +100 | 39 | 50 | 0 | + 50 |
| 34 | 83 | 33 | + 50 | 42 | 100 | 67 | + 33 |
| 56 | 100 | 83 | + 17 | 51 | 100 | 0 | +100 |
| 32 | 80 | 60 | + 20 | 46 | 67 | 50 | + 17 |
| Mean | 75.2 | 35.9 | + 39.3 | Mean | 66.7 | 27.9 | + 38.8 |

negative stimuli as they do to the positive ones. Four of the *Ss* show an equal percentage of responses to both, while three show a slight excess favoring the positive stimulus, three an excess favoring the negative stimulus. By contrast, the *Ss* of Group B, informed in advance about the stimuli, show much more discrimination. Every *S* shows an excess in favor of the positive stimulus equalling or exceeding the greatest discrimination shown by any member of the control group.[3] The initial advantage of Group B, by virtue of instructions, is evident.

In Experiment II, previously reported, ability to state the stimulus-relationships in words accompanied more successful discrimination. Ordinarily, in a conditioned discrimination-experiment, such hypotheses as the *S* is able to adopt within the first few trials will have a relatively low degree of certainty, for there is no assurance from the *S*'s point of view that the one stimulus will continue to be reënforced or that the other will not be reënforced. It may be conjectured that reactions in these experiments are controlled to some extent by the degree of certainty with which particular stimulus-objects are anticipated. The 'hypothesis' given ready-made to Group B by instructions carries the experimenter's assurance that one stimulus will always be positive, the other negative, and the discriminatory reactions of the *Ss* of this group are correspondingly prompt. The *Ss* of Group A at best can only conjecture that the one stimulus will always be reënforced, the other never. The probability of this conjecture becomes greater as more exposures occur agreeing with it. Once the probability is established that the one stimulus will always be positive, the other negative, Group A has potentially as much information as Group B. The gradual increase in the success of discrimination of Group A may correspond to the gradual establishment of the probabilities in the situation.

The curves of Fig. 1 show that the initial advantage of Group B is lost quickly. While there is some residual advantage for Group B as indicated by mean results throughout Day 3, the differences are without significance after the first few trials of Day 2. That the two groups perform essentially alike by trials 25–36 of Day 2 is shown by the data for individual *Ss* presented in Table III. The mean discrimination for the two groups is similar; the median excess of responses to the positive over responses to the negative stimulus is alike (33%) for both groups.

The results thus far presented in Fig. 1 and Tables I-III suffice to indicate the effect of verbal knowledge of the stimuli on the course of conditioned discrimination. Advance information regarding positive and negative stimuli results in much more rapid discrimination, but this initial advantage is soon lost as the non-informed *Ss* become acquainted with the stimuli through repeated exposure to them.

*(3) Success in discrimination and susceptibility to extinction.* Similarities and differences between the various kinds of inhibitory conditioning pose problems important for the theory of conditioning. Two forms of

[3] The difference meets the accepted standards of significance for small samples. For our data, Fisher's $t = 4.0$, whereas $t = 2.9$ is considered satisfactory for a difference based on samples of this size.

so-called internal inhibition are represented in the reported experiment: differential inhibition and experimental extinction. It is possible to compare the success in discrimination with the susceptibility to extinction at the end of the experiment in order to determine the commonality between the two processes. It is evident that discrimination is a process of selective extinction, but this does not deny the possibility of differences between them. Data for the final conditioning trials of Day 3, and the extinction trials which followed, are given in Table IV.

TABLE IV

PERCENTAGE FREQUENCY OF CONDITIONED RESPONSES TO POSITIVE AND NEGATIVE STIMULI WITHIN FINAL CONDITIONING TRIALS (DAY 3, TRIALS 49–60) AND WITHIN EXTINCTION TRIALS (DAY 3, TRIALS 61–70)

(The Ss are arranged in order of frequency of simple conditioning on Day 1.)

| Group | S | Conditioned responses | | | | | |
|---|---|---|---|---|---|---|---|
| | | to positive stimulus | | | to negative stimulus | | |
| | | Final trials of conditioning | Extinction trials | Diff. | Final trials of conditioning | Extinction trials | Diff. |
| A | 44 | 50 | 0 | +50 | 0 | 0 | 0 |
| | 47 | 50 | 33 | +17 | 33 | 17 | +16 |
| | 35 | 50 | 33 | +17 | 0 | 0 | 0 |
| | 29 | 50 | 67 | −17 | 33 | 33 | 0 |
| | 50 | 100 | 33 | +67 | 100 | 83 | +17 |
| | 40 | 50 | 50 | 0 | 0 | 0 | 0 |
| | 52 | 83 | 33 | +50 | 0 | 0 | 0 |
| | 34 | 83 | 67 | +16 | 33 | 0 | +33 |
| | 56 | 100 | 100 | 0 | 100 | 100 | 0 |
| | 32 | 50 | 67 | −17 | 17 | 0 | +17 |
| | Mean | 66.6 | 48.3 | +18.3 | 31.6 | 23.3 | +8.3 |
| B | 49 | 33 | 0 | +33 | 0 | 17 | −17 |
| | 37 | 80 | 33 | +47 | 0 | 0 | 0 |
| | 45 | 50 | 17 | +33 | 17 | 33 | −16 |
| | 43 | 100 | 17 | +83 | 17 | 17 | 0 |
| | 41 | 100 | 17 | +83 | 83 | 67 | +16 |
| | 48 | 83 | 67 | +16 | 40 | 20 | +20 |
| | 39 | 100 | 17 | +83 | 0 | 0 | 0 |
| | 42 | 67 | 67 | 0 | 83 | 67 | +16 |
| | 51 | 83 | 33 | +50 | 0 | 17 | −17 |
| | 46 | 100 | 33 | +67 | 0 | 0 | 0 |
| | Mean | 79.6 | 30.1 | +49.5 | 24.0 | 23.8 | + 0.2 |

Responses to the positive stimulus are treated separately from responses to the negative stimulus. The extinction series differed from the conditioning series only for the positive stimulus, since the negative stimulus was non-reënforced throughout. Responses to the positive stimulus show reduction, as expected, for 15 of the 20 Ss. Responses to the negative

stimulus, already reduced in frequency through the discriminatory conditioning, show little change during extinction. The *Ss* of Group B show somewhat greater mean extinction of responses to the positive stimulus than the *Ss* of Group A, but the difference is not a trustworthy one.[4]

The relation of extinction to success of discrimination is not clear from the data of Table IV. All of the chief possibilities are represented, as follows:

*(1)* Those who discriminate well may also extinquish rapidly. Examples: A-44, A-52, B-37, B-45, B-43, B-39, B-51, B-46.

*(2)* Those who discriminate well may be resistant to extinction. Examples: A-29, A-32, A-34, A-40, B-48.

*(3)* Those who discriminate poorly may nevertheless extinguish rapidly. Examples: A-50, B-41. (It is of interest that in both of these cases extinction is more pronounced in responses to the positive stimulus.)

*(4)* Those who discriminate poorly may also be resistant to extinction. Examples: A-56, B-42.

*(5)* Differences not readily classifiable: A-35, A-47, B-49.

It is not possible to determine from the small population which of the relationships is to be considered typical. While 6 *Ss* of Group B fall into the class of successful discrimination and rapid extinction, as compared with but two members of Group A, the differences are not solely a matter of instructions, for at least one member of each group falls into each of the classes. There are sufficient differences in performances during discriminatory conditioning and extinction that it is unwise to assume that the two processes represent the same underlying mechanism, at least until these differences are accounted for.

*Discussion.* The results confirm the conjecture based on earlier experiments that the capacity to formulate the stimulus-relationships in words facilitates discriminatory conditioning in man. This is but another instance of the importance of verbial controls in human behavior, even in processes supposed to be relatively automatic.

The lack of clear relationship between discrimination and extinction may be based on differences in hypotheses adopted during the brief series of non-reënforced trials. Failure to reënforce the positive stimulus when it was first presented within the extinction series might have been in accord with several possibilities from the *S*'s point of view, such as (1) a slip by the experimenter, whereby reënforcement was inadvertently omitted, (2) a reversal of conditions, so that the stimulus previously

[4] For the difference in amount of extinction between Group A and Group B, Fisher's $t = 2.4$, whereas $t = 2.9$ is necessary for significance as conventionally determined.

positive would become negative, and vice versa, (3) deliberate omission of reënforcement for the whole final series. While the third of these possibilities was the correct one, there was no compulsion upon the *S* to adopt it as his hypothesis until the series was well along. Since the main experiment illustrates the differences which verbal controls make within discrimination, it is not unlikely that such controls also affect extinction selectively. Success in discrimination does not guarantee the adoption of the correct hypothesis within extinction, and therefore there is some lack of correlation between discrimination and extinction.

*Summary.* This study may be summarized as follows:

*(1)* Simple conditioned eyelid responses were established during one experimental period by presenting a visual stimulus followed by an air-puff to one cornea.

*(2)* Discriminatory conditioned eyelid responses were developed during a second and third experimental period by presenting another visual stimulus without reënforcement in random order with the previous stimulus, which was always reënforced.

*(3)* One group of 10 *Ss* (Group A) began the discrimination session without further instructions. Another group of 10 *Ss* (Group B) which had performed similarly to Group A within the simple conditioning trials, was informed before the discrimination trials which stimulus would be positive, which negative. This information resulted in immediate conditioned discrimination by Group B within trials 1–12 of the first discrimination session. Within these trials Group A showed generalization. By trials 25–36, which followed the presentation of 12 positive and 12 negative stimuli, performances of Groups A and B were essentially alike. It is inferred that Group A was self-instructed through exposure to the positive and negative stimuli.

*(4)* Success in discrimination failed to show a clear relationship to susceptibility to extinction in non-reënforced trials at the end of the experiment. While the more successful discriminators tended also to extinguish rapidly, some successful discriminators were resistant to extinction; similarly, some non-discriminators showed prompt extinction, some showed resistance to extinction. It is considered unwise to assume that discrimination and extinction represent a common underlying mechanism until these differences are explained.

*(5)* The results furnish another instance of the significance in man of verbal controls over processes which have been supposed to be relatively automatic.

# VARIABILITY IN THE AFFECTIVE JUDGMENT

By WILLIAM A. HUNT, New York Psychiatric Institute and Connecticut College, and JANE FLANNERY, Brown University

Interest in the judgmental aspects of the affective response has been heightened by Peters's restatement[1] of Carr's judgmental theory of pleasantness and unpleasantness.[2] As one of us has pointed out,[3] the usual affective situation is one in which an individual is called upon to pass judgment upon the pleasantness or unpleasantness of some stimulus. Not only is a bodily affective response involved but there is also a differential, discriminatory reaction which results in a verbal label attached to this response. In the usual experimental setting this verbal response becomes overt and appears as a category of report, *i.e.* the *O* says "pleasant" or "unpleasant." Often times the discriminatory, judgmental processes and the processes of verbal report become the most important and significant parts of the total affective behavior. This change in no way denies the existence of some ultimate, bodily process as a basis for the judgment or discrimination; it merely suggests that many of the phenomena encountered in the field of affectivity may more properly be interpreted as functions of the secondary, judgmental processes which intervene rather than as functions of a primary, specific affective process. Thus, in a recent publication on anchoring effects with an affective scale,[4] Hunt and Volkmann have shown that it is possible to duplicate with the affective judgment the anchoring effects demonstrable in judgments of visual inclination.[5] They suggest the operation of general principles of judgment common to both fields.

In continuing the investigation of the judgmental characteristics of affectivity, we selected variability for study. By *variability* we mean a change in judgment upon repetition of the stimulus. When a stimulus, previously called "very pleasant," is later called "slightly pleasant" or "unpleasant," this shift in the category of report is *variability*. This aspect of the affective response has not drawn much attention in the past, as is shown by its neglect

* Accepted for publication October 19, 1937.

[1] H. N. Peters, The judgmental theory of pleasantness and unpleasantness, *Psychol. Rev.*, 42, 1935, 354-386.

[2] H. A. Carr, *Psychology*, 1925, 290.

[3] W. A. Hunt, The meaning of pleasantness and unpleasantness, this JOURNAL, 45, 1933, 345-348.

[4] Hunt and J. Volkmann, The anchoring of an affective scale, this JOURNAL, 49, 1937, 88-92.

[5] Volkmann, The anchoring of absolute scales, *Psychol. Bull.*, 33, 1936, 742.

in Beebe-Center's recent survey.[6] Usually variability is given a general statistical value or assumed to represent an area within which 'chance' factors are working, and no further attention is paid to it. We propose, however, to view it as a lawful phenomenon, the causal result of significant factors. Our paper studies this variability as a function of the number of stimuli to be judged, the number of categories to be used in judgment, and the position of any particular judgment within the series of judgments made during the experimental session.

In the situation used in our experiment the *Os* were called upon to report the affective value of a series of colors. Standard Milton Bradley colored papers were used as stimuli and were presented through a 6-in. square aperture in a large gray screen. The apparatus is described elsewhere.[7] The *O* sat about 4 ft. from the screen. Lighting was kept constant, and a constant interval between stimuli was preserved. When repeated, each series of colors was presented in the same order, and there was a standard interval between successive series. All of the *Os* were women and were either taking or had taken a course in elementary psychology. They were not informed as to the purpose of the experiment but were led to believe that it concerned the affective value of colors. They were all warned against making memory judgments. Introspections were taken at the end of each experimental period. A preliminary practice session was given each *O* before the start of the experiment in order that she might become familiar with the general procedure.

In the first section of the experiment 6 *Os* were asked to judge 4 different series of stimuli of 3, 6, 9, and 12 colors each. Each series was repeated 5 times during an experimental session, and only one series was judged on any one day. An interval of one week usually intervened between experimental sessions. A scale with 5 categories was used in reporting, and the instructions were as follows: "I am going to present various colors to you. Please judge each one as very pleasant, pleasant, indifferent, unpleasant, or very unpleasant. Do not attempt to make memory judgments, but judge each color as it affects you at the moment of presentation." At a later date, the routine was repeated with 8 new *Os*. The number of repetitions of each series was increased from 5 to 10. There were no other changes in conditions.

In the second section of the experiment 6 *Os* were asked to judge 3 different series of 12 colors each. Each series was repeated 5 times. The instructions were identical with the preceding ones except that for the first series *O* was told to use a scale of 3 categories—*pleasant, indifferent,* and *unpleasant;* while for the second series he was told to us 5 categories—*very pleasant, pleasant, indifferent, unpleasant, very unpleasant;* and for the third series 7 categories—*very pleasant, pleasant, slightly pleasant, indifferent, slightly unpleasant, unpleasant, very unpleasant.* The experiment was again repeated with a new group of 6 *Os,* and the number of repetitions of each series was increased from 5 to 10.

[6] J. G. Beebe-Center, *Pleasantness and Unpleasantness,* 1932, 1-427.

[7] Hunt, The relation of bright and dull pressure to affectivity, this JOURNAL, 43, 1931, 87-92.

Since we have defined variability as a change in the category of report, in scoring we have determined the amount of variability by counting the changes in report for each color during the experimental session. Thus, if a color were called *pleasant* the first two times it was presented, *slightly pleasant* the next two times, and *unpleasant* the last time, there would be two shifts in the category of report and the variability score would be *two*. If a color were called *pleasant* the first time, *slightly pleasant* the second time, *pleasant* the third time, and *slightly pleasant* the fourth and fifth times there would be three shifts in report and the variability score would be *three*. Adding the sums for all the colors in any series gives the total variability for each series.

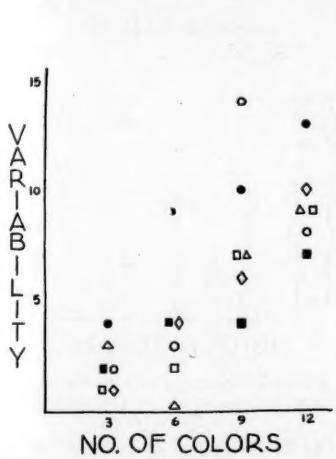


FIG. 1. VARIABILITY OF 6 *Os* WITH 3, 6, 9, AND 12 COLORS REPEATED 5 TIMES

FIG. 2. VARIABILITY OF 8 *Os* WITH 3, 6, 9, AND 12 COLORS REPEATED 10 TIMES

The relation of variability to the number of colors to be judged is shown in Figs. 1 and 2. Fig. 1 is for 6 *Os* with the colors repeated 5 times, while Fig. 2 is for 8 *Os* with the colors repeated 10 times. Each *O* is represented by a different symbol. Variability is shown as ordinates of each graph with the number of colors in the series shown as abscissae. The graphs show that as the number of colors to be judged increases the variability increases. There are large individual differences between *Os* in amount of variability. Considering the probable complexity of the factors involved in the affective judgment, the increase of variability with an increase in the number of colors comes out rather clearly.

The relation of variability to the number of categories to be used in judging the colors is shown in Figs. 3 and 4. Fig. 3 is for 6 *Os* with the colors repeated 5 times. Fig. 4 is for 6 *Os* with the colors repeated 10 times. Variability is shown as ordinates; the number of categories used is shown as abscissae. While the results here are not so clear, there is a tendency for variability to increase with the number of categories used. This difference in the clarity of the increases for number of colors and for number of categories may be due to a difference in the ranges covered, but this question cannot be answered since there is no common basis for comparison. The

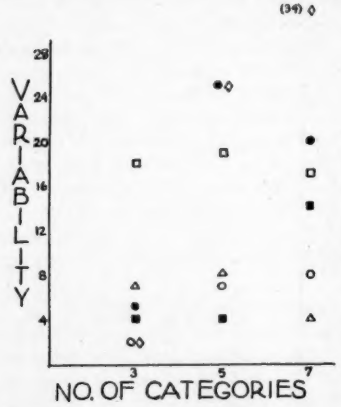


FIG. 3. VARIABILITY OF 6 *Os* USING 3, 5, AND 7 CATEGORIES OF JUDGMENT AND 5 REPETITIONS

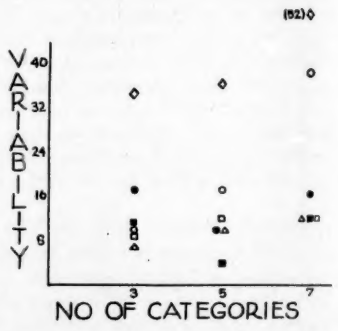


FIG. 4. VARIABILITY OF 6 *Os* USING 3, 5, AND 7 CATEGORIES OF JUDGMENT AND 10 REPETITIONS

relation, in terms of judgment, between the number of things to be judged and the number of categories to be used in judging them is not known. That they should have an equal effect upon the variability of the affective judgment is not to be expected.

That variability should increase with the number of colors and the number of categories is not surprising. The more elements entering into a judgmental situation, the more possibility there is of a person's changing his mind. The experimentally demonstrated fact becomes particularly interesting, however, when one tries to fit it into affective theory. Any theory which attributes affective phenomena exclusively to a process in an end-organ will have difficulty in including this function. When we increase colors and categories we are, to some extent, diminishing the differences

between the adjacent colors and categories. As we add new elements to the situation, we decrease the relative discriminability of any one element involved. This explanation, however, in terms of discrimination and judgment, involves the working of more highly organized levels of response rather than a peripheral process in an end-organ. We need not deny ultimate peripheral involvement of some sort, but our results indicate the importance of the judgmental nature of the affective response.

The relation of variability to the position of the series in the experimental period is of interest. We have noted the amount of variability appearing during the successive repetitions of the series. We have taken the judgments for each series of colors and compared the total variability shown on all the colors in the series with the total variability shown in later repetitions of the series. Since the data are uniform in tendency throughout the entire experiment we need exhibit only two typical examples. Table I shows the results for 14 *O*s judging a series of 12 colors repeated 10 times, and using 5 categories. The table shows the number of changes of judgment (variability) for each *O* upon each repetition of the series. In general, variability decreases as the number of repetitions increases. As we have said, this decrease occurs, not only with these two groups of cases, but with all our data. The trend is uniform throughout the entire experiment. As the individual experimental session progresses, variability tends to disappear and the affective judgment tends to stabilize. In some cases there is a preliminary rise during the first few presentations, which seems to be due

TABLE I

NUMBER OF CHANGES IN JUDGMENT (VARIABILITY) DURING SUCCESSIVE PRESENTATIONS OF THE COLOR SERIES

| | Presentation | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| *O* | 2nd | 3rd | 4th | 5th | 6th | 7th | 8th | 9th | 10th |
| A | 2 | 0 | 2 | 3 | 0 | 0 | 1 | 0 | 0 |
| B | 2 | 3 | 1 | 3 | 1 | 1 | 2 | 0 | 1 |
| C | 1 | 3 | 2 | 1 | 1 | 1 | 1 | 1 | 1 |
| D | 5 | 2 | 2 | 1 | 0 | 1 | 1 | 1 | 4 |
| E | 2 | 0 | 1 | 2 | 4 | 2 | 3 | 2 | 3 |
| F | 1 | 2 | 1 | 4 | 2 | 4 | 4 | 4 | 3 |
| G | 8 | 4 | 4 | 2 | 1 | 2 | 3 | 1 | 2 |
| H | 5 | 7 | 2 | 0 | 5 | 8 | 7 | 4 | 5 |
| I | 4 | 5 | 5 | 1 | 2 | 4 | 1 | 1 | 2 |
| J | 5 | 2 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| K | 4 | 3 | 5 | 2 | 0 | 3 | 3 | 3 | 2 |
| L | 2 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| M | 3 | 2 | 1 | 1 | 0 | 0 | 0 | 0 | 0 |
| N | 4 | 2 | 2 | 1 | 2 | 2 | 4 | 0 | 2 |

to the fact that the contrast effects make the affective value of each color a function of the other colors within the series in which it appears. Only when this series has been established (comprehended by the *O*) can a final, definite affective value be assigned to any individual color. Sometimes the series is established in one presentation, sometimes not. These latter instances in which the series is not immediately established in one presentation appear to explain the cases in which there is a brief rise in variability before the drop begins. Moreover, this interpretation is borne out by the introspections of the *O*s. Once the series is established, however, the variability decreases.

What some of these other processes may be becomes clearer upon further inspection of the data. If one examines these curves of decreasing variability with no thought for the fact that they are concerned with the 'variability' of the 'affective judgment,' he may immediately be impressed by their similarity to learning curves. Variability functions like error in a learning situation. Nor is this analogy superficial. To say that, once the series of colors is established and affective values are assigned to it, variability gradually disappears, is merely another way of saying that once the *O* has decided upon the affective value of the colors he proceeds to learn the responses, with variability representing the error appearing in this learning process. This learning takes place despite the fact that the *O*s have been specifically warned against making memory judgments.

If we take another aspect of the learning situation, time, and investigate it in the affective situation, again we find a similarity. As learning progresses, the speed of response increases. This relation also holds for the affective judgment. Unpublished data obtained by the senior author with the assistance of Miss Elizabeth Johnson show that, as color series are repeated, the affective judgments become more rapid. The curve of decreasing reaction-time parallels closely the curve of decreasing variability. After a number of repetitions a basic response-time is reached and held, provided, of course that adequate preliminary sessions have already removed the general practice effect. Moreover, such a decrease in response-time is specific for the particular series of colors being judged. If, once the basic response-time has been reached, the color stimuli are changed and the *O* is given a new series to judge, the reaction-time again lengthens, decreasing later with the subsequent repetition of this new series until the previous base-line is again approximated. Since general practice effect has been taken care of by repeated training in the general situation, this specific practice effect appearing within each new series bears a strong resemblance to the decrease in performance-time typical of learning.

There is a further argument for our claim that the judgmental situation is also a learning situation. Examination of the data shows that the decrease in variability is not an uninterrupted one. There are occasional temporary increases in variability, causing humps in the descending curve. Although these accelerations might have been disregarded as 'chance' factors, they are easily understood if the *O*s are questioned. The introspective reports show that as the *O* continues repeating his judgments, there often comes a time when he realizes that the judgment has become a formal, habitual reaction and does not represent his true affective response to the color. One *O* said: "I suddenly realized that my judgments didn't really represent the way I actually felt about the colors, so I changed them." She had not been conscious of making memory judgments but was actually doing so. She learned a verbal response which later (as she suddenly realized) no longer represented her true affective attitude.

Our claim for the similarity between the affective judgment and the learning process is also borne out by our original finding that variability increases with an increase in the number of things to be judged and the number of categories to be judged and the number of categories to be used in judging them. In learning one finds error increasing with the number of things to be learned.

While the above introspections point out the importance of learning as a factor in the affective judgment, they also show that learning is not the only factor involved. There seems to be some basic factor (and we favor a view that definite somatic cues are responsible) on the basis of which affective judgments are rendered. Once judgment has taken place, however, learning enters in and a fixed, habitual response may result which often, as our work shows, departs from the real affective value of the stimulus in question. In other words, we have here experimental proof of our earlier suggestion that the affective judgment may represent nothing but a learned verbal response. Many of the classical problems in affectivity will be clarified if this is kept in mind.

## Summary

Affective variability is defined as a change in the category of report used in judging the affective value of a stimulus. In judging colors, this variability is shown to increase both with an increase in the number of colors to be judged and an increase in the number of categories to be used in judging them. Within any single series of colors repetition results in a reduction of variability. These phenomena are interpreted as indicating the participation of judgment and learning in the affective response.

# SOME FACTORS IN SIZE-CONSTANCY

By NORMAN M. LOCKE, Hunter College

The basic determinants of the size of an object are its physical dimensions and its distance from the eye. The size of the image on the retina corresponds to the direct ratio of height to distance. With a constant height at various distances this ratio has various values and the size of the retinal image likewise changes. The phenomenon of size constancy is met when the individual judges the height of the object apart from the physical projection on the retina. A rather correct judgment of size is made when the image on the retina indicates a widely different response. In other words, the judgment does not follow the stimulus-value; somehow or another the individual corrects for the variation in distance.

The mechanism of this correction, with respect to the individual, may be either internal or external or both. The factors which have been thought to make for size-constancy will be briefly indicated.

*Kinesthetic factors.* Because of the adjustment of the eye made in focussing on an object and the consequent energy change, Frank states that the tension of the musculature involved in accommodation and convergence is, to a slight degree, a cue to constancy.[1] She found that a square at a fixed distance appeared larger when it was directly fixated than when it was behind the point of fixation, and that it appeared smaller when it was in front of it. Hermans agrees with this explanation, stressing the kinesthetic cues involved.[2] In a study of the constancy of the after-image, Helson concluded that accommodation was the chief factor which made for the phenomenon.[3] Weber and Bicknell found size-constancy judgments changing with an increase in disparation.[4] The increase caused the area of the object to appear smaller, with a consequent increase appearing in the tendency to correct for size. Holaday studied twenty-eight conditions, both internal and external, and reported a wide range of constancy scores.[5] These conditions were such factors as monocular judgment, the introduction of head-movements, the use of a reduction tube, squinting, a differing fixation-point, looking through a camera, and various sets.

*Stimulus-factors.* Some further conditions in Holaday's study included different

* Accepted for publication October 15, 1937.

[1] H. Frank, Über den Einfluss inadäquater Konvergenz und Akkommodation auf die Sehgrösse, *Psychol. Forsch.*, 13, 1930, 135-144.

[2] T. G. Hermans, Visual size constancy as a function of convergence, *J. Exper. Psychol.*, 21, 1937, 145-161.

[3] H. Helson, Size-constancy of the projected after-image, this JOURNAL, 48, 1936, 638-642.

[4] C. O. Weber, and N. Bicknell, The size-constancy phenomenon in stereoscopic space, this JOURNAL, 47, 1935, 436-448.

[5] B. E. Holaday, Die Grössenkonstanz der Sehdinge bei Variation der inneren und äusseren Wahrnehmungsbedingungen, *Arch. ges. Psychol.*, 88, 1933, 419-486.

changes in illumination. He also used a milk-glass plate illuminated from the rear as the test object, and different sized objects. Beyrl found that three-dimensional objects make for higher constancy than do two-dimensional objects.[6] There was greater constancy found when boxes were used as the test objects than when discs were used. Thouless[7] and Smirnov and Volokitina[8] report constancy scores taken at different distances. Higher constancy is usually found when the objects are nearer the observer than when the same objects are further away. Frank[9] and Katona[10] discovered that the level of the test-object with respect to the eye is a factor in making the size-judgment. The nearer the eye to the level of the object the lower the resulting score. Smirnov emphasized the angle of the line of fixation as a factor in size-constancy.[11] In the study of Helson mentioned above the factors of "covering" and clearness are included as explanations of size-constancy.

*Methodological factors.* Frank,[12] Grabke,[13] Katona,[14] and Martius[15] found that simultaneous perception of the standard and variable lowers size-constancy. Burzlaff reported differences in judgment resulting from different methods of testing.[16] When his *Ss* saw the entire range of the variable stimuli they made higher scores than when one variable size at a time was presented.

Since these judgments are seemingly made with respect to the constellated visual field, it was felt that the introduction of a constant reference point would demonstrate the influence of that factor. The presence of shadows in a visual field produces interesting shifts in the resulting perception. It was decided to obtain size-constancy judgments based on a simple stimulus-situation and then to vary the appearance of the field by differential shading.

## Method and Procedure

The stimulus-objects were gray cardboard squares, facing the *Ss* in the frontal

[6] F. Beyrl, Über die Grössenauffassung bei Kindern, *Zsch. Psychol.*, 100, 1926, 344-371.

[7] R. H. Thouless, Phenomenal regression to the real object, *Brit. J. Psychol.*, 21, 1931, 339-359.

[8] A. A. Smirnov and M. M. Volokitina (Dependence of the constancy of size perception upon the relative intervals between objects at different distances from the observer), in *Visual sensation and perception*, Vol. II, ed. by S. V. Kravkov and B. M. Teplov, 1935.

[9] H. Frank, Die Sehgrössenkonstanz bei Kindern, *Psychol. Forsch.*, 10, 1928, 102-106.

[10] G. Katona, Experimente über die Grössenkonstanz, *Zsch. Psychol.*, 97, 1925, 215-252.

[11] A. A. Smirnov, (Dependence of the constancy of size perception of objects upon distance and the angle with the fixation line.), in *Visual sensation and perception*, Vol. II, ed. by S. V. Kravkov and B. M. Teplov, 1935.

[12] H. Frank, *op. cit.*

[13] H. Grabke, Über die Grösse der Sehdinge im binokularen Sehraum bei ihrem Auftreten im Zusammenhang miteinander, *Arch. ges. Psychol.*, 47, 1924, 237-301.

[14] G. Katona, *op. cit.*

[15] G. Martius, Über die scheinbare Grösse der Gegenstände und ihre Beziehung zur Grösse der Netzhautbilder, *Philos. Stud.*, 5, 1889, 601-617.

[16] W. Burzlaff, Methodologische Beiträge zum Problem der Farbenkonstanz, *Zsch. Psychol.*, 119, 1931, 177-235.

parallel plane, held upright by blocks glued to the back. The standard object was the smaller one, 15 cm. square, and was always placed 50 cm. from the front of the table on which both objects stood. The comparison was a 16½ cm. square. Both standard and comparison were shifted from the right to the left sides of the table at irregular intervals. The table was 489 cm. long, 50 cm. wide, and 80 cm. high, painted flat black, with meter sticks attached to both outside edges. The outfield was restricted by large gray painted beaverboard screens which extended around three sides of the table and almost reached the ceiling. Three 150-w. mazda lamps placed 1 m. apart were suspended from the ceiling and directed in such a way that there were no shadows on the table.

TABLE I

AVERAGE DISTANCE OF JUDGMENT; SIZE OF RETINAL IMAGE FOR *E* (EQUATION SIZE), AND *P* (PROJECTION SIZE); AND RESULTING *I*-RATIO (INTENTION-RATIO) FOR CONDITION A—*i.e.* CLEAR FIELD

*St* (standard size)=0.075

| *S* | Distance | *E*-size | *P*-size | *I*-ratio |
|---|---|---|---|---|
| 1 | 290.0 | .0569 | .0517 | .7426 |
| 2 | 320.0 | .0516 | .0469 | .7970 |
| 3 | 290.0 | .0569 | .0517 | .7426 |
| 4 | 360.0 | .0458 | .0417 | .8400 |
| 5 | 277.5 | .0595 | .0541 | .7089 |
| 6 | 312.5 | .0528 | .0480 | .7865 |
| 7 | 252.5 | .0653 | .0594 | .5943 |
| 8 | 305.0 | .0541 | .0492 | .7750 |
| 9 | 350.0 | .0471 | .0429 | .8331 |
| 10 | 340.0 | .0485 | .0441 | .8210 |
| 11 | 320.0 | .0516 | .0469 | .7970 |
| 12 | 345.0 | .0478 | .0435 | .8270 |
| 13 | 360.0 | .0458 | .0417 | .8400 |
| 14 | 347.5 | .0475 | .0432 | .8280 |
| 15 | 350.0 | .0471 | .0429 | .8331 |
| 16 | 315.0 | .0552 | .0476 | .6744 |
| 17 | 367.5 | .0449 | .0408 | .8430 |
| M | 323.7 | .0510 | .0463 | .7957 |

The *Ss* were college students, girls, average age about 18 yr. They sat 150 cm. from the end of the table. They were prevented from seeing the field between trials by a white cloth screen.

There were four stimulus-conditions. The procedure was the same in each. The *Ss* were instructed to report which of the two objects appeared larger—not the one that was larger, nor the one they thought should be larger, but just how they looked. The screen was drawn, permitting a 3-sec. exposure of the visual field, and *S* reported. The comparison object was moved back, away from *S*, in 5-cm. steps until *S* consistently reported the physically smaller object as larger. A reading was taken of the distance of the variable from *S* by glancing at the meter sticks on the edge of the table. The method of limits was used in presenting the stimuli, the *Ss* being given 4 ascending and 4 descending series.

Series A was the normal basic condition in which the *Ss* saw the table and the test-objects. In Series B two posts were introduced, which served as reference points.

TABLE II

AVERAGE DISTANCE OF JUDGMENT; SIZE OF RETINAL IMAGE FOR *E*, AND *P*; AND RESULTING *I*-RATIO FOR CONDITION B—*i.e.* REFERENCE POSTS PRESENT

($St$=0.075)

| S | Distance | *E*-size | *P*-size | *I*-ratio |
|---|---|---|---|---|
| 1 | 305.0 | .0541 | .0492 | .7750 |
| 2 | 360.0 | .0458 | .0417 | .8400 |
| 3 | 317.5 | .0520 | .0472 | .7908 |
| 4 | 355.0 | .0465 | .0423 | .8344 |
| 5 | 325.0 | .0508 | .0462 | .8038 |
| 6 | 345.0 | .0478 | .0435 | .8271 |
| 13 | 460.0 | .0359 | .0326 | .8842 |
| 14 | 332.5 | .0496 | .0451 | .8130 |
| 15 | 375.0 | .0440 | .0400 | .8484 |
| 16 | 307.5 | .0537 | .0488 | .7772 |
| 17 | 415.0 | .0398 | .0361 | .8665 |
| M | 354.3 | .0466 | .0423 | .8308 |

These posts were wooden columns, 2 cm. square, mounted on a base of 5 cm. width, the entire post measuring 25 cm. high. They were painted gray and were placed in the same plane and to the side of the small stimulus-object. In Series C a 150-w. lamp was directed at the table from out of the field and the other lights were turned out. It was arranged so that the light fell to the rear of the table illuminating the large test-object. The smaller standard in front was in shadow. The standard object matched paper no. 40 of the Hering gray series, and the variable object matched paper no. 30. In Series D the standard front square was in the light, matching Hering paper no. 12 and the shaded rear card matched paper no. 35.

## RESULTS

Tables I-IV present the mean distance at which the *S*s reported the physically larger square to be smaller than the physically smaller square. It will

TABLE III

AVERAGE DISTANCE OF JUDGMENT; SIZE OF RETINAL IMAGE FOR *E*, AND *P*; AND RESULTING *I*-RATIO FOR CONDITION C—*i.e.* FRONT FIELD SHADED

($St$=0.075)

| S | Distance | *E*-size | *P*-size | *I*-ratio |
|---|---|---|---|---|
| 1 | 317.5 | .0520 | .0472 | .7908 |
| 2 | 335.0 | .0493 | .0448 | .8146 |
| 3 | 360.0 | .0458 | .0417 | .8400 |
| 4 | 387.5 | .0426 | .0387 | .8549 |
| 5 | 355.0 | .0465 | .0423 | .8344 |
| 6 | 357.5 | .0462 | .0420 | .8356 |
| 9 | 410.0 | .0402 | .0366 | .8694 |
| 10 | 600.0 | .0275 | .0250 | .9132 |
| 11 | 385.0 | .0429 | .0390 | .8542 |
| 12 | 435.0 | .0379 | .0345 | .8790 |
| M | 394.3 | .0418 | .0380 | .8598 |

be remembered that the size of the retinal image is determined by the height of the object divided by the distance of the object from the eye. The physically smaller object was 15 cm. high and stood 200 cm. from *S*. The retinal image of this standard, which shall be termed *St* was therefore 0.075. The physically larger object was 16½ cm. high and the distance varied with every *S*, as given in the tables. If we divide this height by the several judgment distances we obtain a ratio for these retinal images. This

TABLE IV

AVERAGE DISTANCE OF JUDGMENT; SIZE OF RETINAL IMAGE FOR *E*, AND *P*; AND RESULTING *I*-RATIO FOR CONDITION D—*i.e.* REAR FIELD SHADED

(St=0.075)

| S | Distance | E-size | P-size | I-ratio |
|---|---|---|---|---|
| 1 | 277.5 | .0595 | .0541 | .7089 |
| 2 | 285.0 | .0579 | .0526 | .7394 |
| 3 | 295.0 | .0559 | .0508 | .7547 |
| 4 | 370.0 | .0446 | .0405 | .8438 |
| 5 | 270.0 | .0611 | .0556 | .6854 |
| 6 | 295.0 | .0559 | .0508 | .7547 |
| 9 | 322.5 | .0512 | .0465 | .7987 |
| 10 | 315.0 | .0524 | .0476 | .7889 |
| 11 | 305.0 | .0541 | .0592 | .7750 |
| 12 | 287.5 | .0574 | .0522 | .7382 |
| M | 302.3 | .0546 | .0496 | .7678 |

will be termed the equation size and designated as *E*. It can readily be seen that in all instances *St* is larger than *E*. This means that the retinal image of the physically smaller object was larger than that of the physically larger object, yet the *S*s persisted in reporting the retinally smaller object as the phenomenally larger one. In short, there was a fairly high degree of size-constancy.

The degree of size-constancy can be determined by taking into account the hypothetical size of the standard if it were projected to the distance obtained by the judgment of the comparison. The consequent ratio of this projection-size *P* with *E* and *St* yields a measure of constancy. The equation-point *E* can range between *P* and *St* with a resulting shift in the percentage of constancy. A formula to express this is $(\log St - \log E)/(\log St - \log P) = I$ or intention-ratio. A more complete discussion of this treatment of the data along with an exposition on the optics involved is given by Locke.[17]

The average *I*-ratio for Series A, the normal clear field, is 0.80. This average is raised to 0.83 if reference posts are in the field of view when

[17] N. M. Locke, A comparative study of size constancy, *Ped. Sem.*, 51, 1937, 255-265.

making the size-constancy judgment. When the nearer standard object is shaded the *I*-ratio increases to 0.86, and when the farther comparison object is shaded it drops to 0.77.

By using the average judgment distances we can discover whether the differences between one situation and the others are statistically reliable. Because of the small number of cases in each pair, Fisher's method of testing the significance of mean differences was used.[18] In this instance the

TABLE V

MEANS AND SIGNIFICANCE OF MEAN DIFFERENCES FOR ALL CONDITIONS, USING THE SHORT SAMPLE METHOD

| Condition | $M_1$ | $M_2$ | N | t | P |
|---|---|---|---|---|---|
| Clear-Posts | 326.4 | 354.3 | 11 | 4.93 | .0100 |
| Clear-Front | 320.5 | 394.3 | 10 | 3.39 | .0100 |
| Clear-Rear | 320.5 | 302.3 | 10 | 2.94 | .0172 |
| Clear-Posts | 308.3 | 339.1 | 6 | 4.00 | .0100 |
| Clear-Front | 308.3 | 352.1 | 6 | 4.25 | .0100 |
| Clear-Rear | 308.3 | 298.7 | 6 | 1.45 | .2081 |
| Posts-Front | 339.1 | 352.1 | 6 | 1.33 | .2456 |
| Posts-Rear | 339.1 | 298.7 | 6 | 3.15 | .0227 |
| Front-Rear | 352.1 | 298.7 | 6 | 5.62 | .0100 |

scores were taken as paired and assumed to be correlated. In Table V the various differing conditions are shown tested for reliability. The usual critical ratio of 3 changes with the number of cases in Fisher's *t* distribution. He feels that any difference yielding a *P* of 0.05 or less is a reliable difference. This would mean that the probability was 95 in 100 that the obtained difference was a true difference. It will be seen that the difference in equation point made with a clear field and that made when there were reference posts is a reliable one, the probability being 9999 chances in 10,000. The presence of posts makes for higher constancy. The difference between judgments made with a clear field and a front-shaded field is reliably in favor of the latter. In one sample the chances are 9828 in 10,000 that a clear field produces a higher score than a rear-shaded field, and in the other the difference is not reliable. Making cross comparisons, the condition of posts in the field of view and that of the front shadow is not reliably different, but both of these are reliably greater than the rear-shaded field.

One would expect that the presence of reference points in a size-constancy judgment would increase the degree of that constancy. The cue presented definitely assists in the perception of depth and helps maintain

[18] R. A. Fisher, *Statistical Methods for Research Workers*, 1936.

the stability of perceptual judgment despite the opposed stimulus-value. Presumably, the more posts in the visual field the greater the size-constancy, within reasonable limits.

When the front part of the field was thrown into shadow the standard in front appeared larger, and consequently the resulting equation was made at a distance greater than that of the clear field. When the rear part of the field was shaded and the front illuminated, the small test-object appeared smaller with a consequent shortening of the equation distance. This difference was not as great as in the opposite situation, doubtless because of the compensating tendency of the shaded comparison object to appear to increase in size. One *S* reported that "the light emphasizes the outline of the smaller, and the larger is blurred against the background."

## Summary and Conclusions

*(1)* Four stimulus-conditions were devised to determine size-constancy values. A fairly high degree of constancy was found for all conditions.

*(2)* The presence of posts as reference points in the visual field increased the degree of size-constancy.

*(3)* When the standard test-object was shaded and the comparison was illuminated the resulting equation was higher than found with a normally illuminated field.

*(4)* When the standard test-object was illuminated and the comparison was shaded the equation was lower than that of the normally illuminated field.

# ON THE PITCH OF FREQUENCY-MODULATED TONES

By R. E. P. YOUTZ, Columbia University, and S. S. STEVENS, Harvard University

A tone whose frequency varies continuously between two limits produces an effect known as a vibrato. The importance of the vibrato in music has led to many studies of its subjective effects—particularly its effects upon the pitch of a note. In a summary of the early experimental work on this problem there is the statement that the vibrato is heard as one salient pitch very near the mean pitch of the oscillations in the vibrato tone, and that, when the range of the vibrato is wide, the pitch is less accurately determined.[1]

From theoretical considerations, Stewart has suggested that the effect of a frequency-modulation, or vibrato, would be to make the pitch of a note uncertain.[2] This conclusion he derived from the assumption that Heisenberg's principle of uncertainty has its analogue in acoustics. This principle applies to any type of periodic phenomenon, and it states that the accuracy with which the frequency of a periodic phenomenon can be determined is proportional to its duration. In other words, in the case of a sound wave, the shorter the duration, the more uncertain is the frequency. Stewart assumed that this rule could be extended to cover the case of the vibrato, because there the duration of any particular small band of frequencies is very short, and the pitch, consequently, should be uncertain.

There can be no doubt about the fact—the pitch of a vibrato is less accurately determined than that of a steady tone. Nevertheless, the explantation which would attribute this fact to the continuous change occurring in the frequency of the tone must be examined more closely. The ear contains a set of resonators, each tuned to respond maximally to a different frequency, and a vibrato is really a group of different frequencies. Consideration of these two additional facts opens to us an entirely new approach to the problem of the vibrato.

It can be shown, both mathematically[3] and physically, that a frequency-modulated tone consists of a number of constant, non-modulated fre-

* Accepted for publication September 13, 1937.

[1] J. Tiffin and H. Seashore, Summary of the established facts in experimental studies on the vibrato up to 1932, *Univ. Iowa Stud., Stud. Psychol. Music.* 1, 1932, 344-382.

[2] G. W. Stewart, Problems suggested by an uncertainty principle in acoustics, *J. Acoust. Soc. Amer.*, 1, 1931, 325-329.

[3] A. Gray, G. B. Matthews and T. M. MacRobert, *Treatise on Bessel Functions*, 1922, 323.

quencies existing simultaneously. The magnitude and frequency of these components depend upon the rate and range of the modulation. Thus, a tone that is varying 8 times per sec. from 989 to 1011 ~ and back again may be analyzed into three principal component frequencies: 992, 1000 and 1008 ~. Here the rate is 8 ~ (sinusoidal variation) and the range is 22 ~. The relative amplitudes of the components under these conditions are as shown in Fig. 1. Other components, beyond those shown and spaced 8 ~ apart, are present, but their amplitudes are much smaller.

Now, if the ear is tuned along the length of the cochlea, the basilar membrane must be stimulated at points corresponding to the different

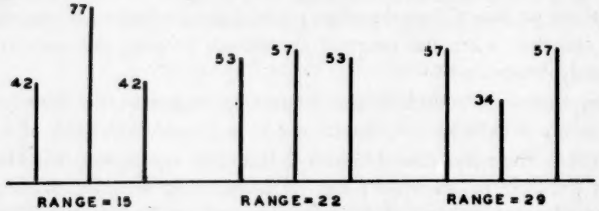


FIG. 1. THE RELATIVE AMPLITUDES OF THE THREE PRINCIPAL COMPONENTS FOR THREE RANGES OF FREQUENCY-MODULATION

In each case the rate of modulation is 8 per sec. so that the components are spaced 8 cycles apart. These amplitudes were computed mathematically and also verified by measurement with a sharply tuned wave-analyzer.

components present in the modulated tone. The tuning is not sharp, however, and excitation due to one component overlaps that due to the others. As a result, beats are produced, alternate reënforcements and cancellations, which give to the vibrato its throbbing character. Nevertheless, if the components are sufficiently separated, the maxima of the excitation due to each may stand out enough for an observer to hear the components separately.

To test this possibility we used the three patterns of modulated tones shown in Fig. 1. These tones were selected so as to make the central components (1000 ~), larger, about equal, and smaller than the two adjacent side-bands.

*Apparatus and procedure.*—The modulated tone was produced by a General Radio Beat-Frequency Oscillator, Type 613, modified by the addition of a rotary condenser in the plate circuit, so that tones with modulation-ranges of 15, 22, and 29 ~ could be produced. The condenser was shaped to give a sinusoidal modulation. A constant tone was generated by a General Radio Oscillator, Type 713-A. This oscillator was fitted with a large scale, calibrated in units of 100 ~, and with a smaller increment scale by which the pitch of the tone could be increased or de-

creased 50 ~. The smaller scale was calibrated in 1 ~ units, so that, in reading it, 1/10 ~ could be estimated. The tones from the two oscillators were put through a relatively clickless 'fader' switch and were presented alternately for 2 sec. each by an automatic relay switch. Each tone was presented to *O* in a pair of crystal earphones. All observations were made in a sound-shielded room.

Three *O*s, chosen for acuity in pitch-discrimination, were asked to set the constant tone to the pitch which they heard in the modulated tone. Before each setting, the constant tone was placed at 30 to 50 ~ above or below the central component of the modulated tone. The *A*bove-*B*elow order was *ABBAABABBA*.

On the first day, after becoming familiar with the apparatus, reading the instructions, and making 5 practice trials, *O* made 10 settings under each of the following conditions: (1) with the range of the modulated tone at 15 ~, (2) at 22 ~, and (3) at 29 ~. Thirty settings were also made on each of the two following experimental days, but the sequence of modulation-ranges was changed on the second day to 22, 29, and 15 ~, and on the third day to 29, 15, and 22 ~, in order to minimize practice and fatigue effects.

The *O*s were allowed to turn the dial controlling the constant tone in one direction only during any one setting. This limitation was imposed on the hypothesis that, if hunting were prevented, the less sophisticated of the *O*s might set the constant tone to the first component of the modulated tone where a subjective equality was to be found, *i.e.* one of the side-bands. The more sophisticated *O*s, if they overshot the central component, might match a farther component. There would then be a clustering of the judgments about the side-band frequencies, as well as around the central frequency.

*Results.* The data give occasional indications of a grouping of the settings about frequencies other than the central frequency. Fig. 1 shows the relative strength of the three principal components for vibratos of 8 ~ with modulation-ranges of 15, 22, and 29 ~. Fig. 2 shows the number of settings actually made by all 3 *O*s at different frequencies in the three modulation-ranges. We find an increasingly wide distribution of settings with increasing range of modulation. Some of the individual records and the verbal reports suggest that this result is to be explained in terms of the position and magnitude of the several components of the modulated tone.

All the *O*s reported trouble in setting the constant to the modulated tone when the range of the modulated tone was at 29 ~. One accused the experimenter of "blotting out the equivalence areas." Another was able to set the constant tone to the upper component of the modulated tone with a very fair degree of accuracy. (Means were +8.99, +7.07, and +7.91.) In one other case the ten settings were grouped about the lower component. (The mean was —7.77.) A plot of the distributions represented by these means indicates that the groupings are definitely around the side-band and are not a chance scatter.

When asked for verbal descriptions of what they had done, all 3 *O*s reported somewhat as follows concerning the ranges of 22 and 29 ~. When the constant tone was below the modulated tone and was being moved up toward a pitch heard in the modulated tone, there would appear a point where equivalence was almost reached, but with further adjustment of the constant tone the equivalence, which should have been at-

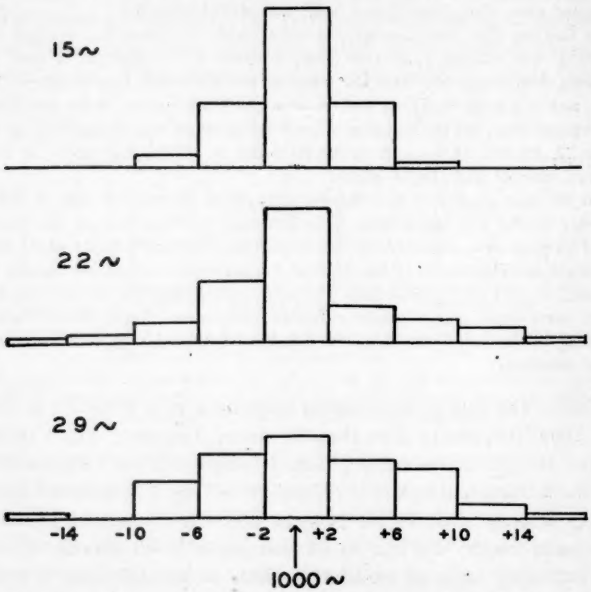


FIG. 2. DISTRIBUTIONS OF SETTINGS FOR EACH OF THE THREE RANGES OF MODULATION
A total of ninety settings was made at each range of modulation.

tained, would vanish, and the dominant pitch in the modulated tone would then appear to be much higher than the constant tone. When *O* had moved the constant tone up to the new, higher pitch in the vibrato, the upper component would then disappear and the dominant pitch would seem to be lower down, back at its first position. (*E* also observed this repeatedly, when acting as *O*.)

The converse phenomenon was observed when the constant tone was started above the modulated tone and gradually moved down to it. The upper component of the modulated tone was clearly heard until the con-

stant tone reached it. Then it disappeared and the lower component (side-band) was heard. Yet when the constant tone was moved down to the lower component, the lower component disappeared and the upper appeared again. (Several *O*s regarded this phenomenon as a trick of *E*'s, but were assured that it was not.) Since this phenomenon was very generally reported, it was regarded as valid and as affording a possible clue to further procedures in identifying subjectively the components in the 29 ~ modulated tone. The subjective dominance of the upper component in the modulated tone, when the constant tone was at the lower component (and the converse situation), was thought possibly to be the result of one or both of these two factors:

*(1) Fatigue of a section of the basilar membrane under continuous stimulation.* Fatigue might occur, for instance, when the constant tone was at 992 (1000–8) ~. Since the constant and modulated tones were heard alternately with no appreciable pause between the cessation of one tone and the beginning of the other, the portion of basilar membrane corresponding to a frequency of 992 ~, being under continuous stimulation, might have its threshold raised slightly, so that the other side-hand would stand out as the dominant one when the modulated tone was heard. A similar fatigue effect would also be found when the constant tone was at 1008 ~.

*(2) A figure-ground effect.* When the constant tone was at 992 ~, it was continuous with the 992-component in the modulated tone, providing a background of continuous stimulation at 992 ~, upon which the 'figure' of the other component at 1008 ~ appeared during every alternate 2 sec. A similar situation arose when the constant tone was at 1008 ~ and 992 ~ appeared as the 'figure.'

On the 'fatigue' hypothesis *(1)*, if an *O* listened to two tones at 1000 and 992 ~ for 15 min. and then suddenly switched to the modulated tone, the 1008 ~ component should stand out sharply, and he should then be able to set the constant tone to the 1008-frequency for as long as the 'fatigue' lasted. This method was tried a number of times by the *E*, on himself and on two others. No positive results were obtained. If the diminished sensitivity occurred, its duration was not great enough to allow settings to be made.

In testing the 'figure-ground' hypothesis *(2)*, the following procedure was used. *O* listened alternately to the modulated tone and to a constant tone of 992 ~. When 10-15 pairs had been heard, *O* was given the constant tone and asked to set it to the pitch at which he had heard the modulated tone. This procedure was most successful when the constant

tone was at a low intensity (6 db. less than the modulated tone). Several such series of settings were made by *E.* (Other *O*s who had not listened to the tones for any length of time found it too difficult to do.) The major finding was that, when the constant tone was set at 992 ~ and paired with the modulated tone, the later settings of the constant tone were always above 1000 ~, showing that the subjective pitch of the modulated tone was higher. However, the means of these settings did not approximate very closely the frequency of the upper side-band of the modulated tone.

We cannot, as yet, be certain of the explanation of the elusive behavior of the components in a modulated tone—why they appear to vanish when the constant tone is set equal to them—but we can be sure of their presence and of the fact that they can be distinguished subjectively under the proper conditions.

*Conclusions.* The subjective effects of a frequency-modulated tone can best be understood by considering the tone to be a set of steady components all existing simultaneously. A single one of these components can be distinguished by the listener under the proper conditions. The pitch of a vibrato becomes more indeterminate as the range of modulation increases, but this change appears to be due to the fact that, as the range increases, the relative amplitude of the outer side bands increases. In other words, a vibrato of wide range excites a more extensive area of the basilar membrane.

The principle of uncertainty does not apply to the case of frequency-modulated tones as has been suggested, because the components of such tones are steady and continuous, and the problem of short duration therefore does not arise. The entire situation can best be conceived in terms of a set of fixed frequencies acting on a set of tuned resonators whose tuning is not very sharp.

# EFFECT OF PRAISE AND REPROOF UPON MEMORY SPAN PERFORMANCE

By ALBERT B. BLANKENSHIP, The Psychological Corporation, New York, and JOHN F. HUMES, Pennsylvania State College

Memory span is a widely used test in clinical psychology. Leaming has defined memory span as "the number of discrete units over which an individual can successively distribute his attention and still organize them into a working unit."[1] To put it more simply, memory span is the ability of an individual to reproduce immediately, after one presentation, a series of discrete stimuli in their original order.

It has been shown that there are a large number of variables which affect memory span performance. These include, among others, the type of material used, the rate of presentation of the stimuli, the fatigue of the subject, distraction, and practice.[2] Since the test employed is so extensively used in clinical practice, it is essential for the clinician to control as many of these variables as possible.

The effect of praise and reproof upon memory span performance has not been thus far reported in the literature. It would seem, however, that these factors are classificatory and therefore might well be investigated. Experiments have already been performed to show that praise and reproof do affect other types of activity. In motor tests, for example, the effects of praise and reproof have been studied. Gates and Rissland,[3] using 74 Barnard College students as *Ss*, found a slight difference in performance in favor of encouragement or discouragement rather than mere repetition of "The Three-Hole Test," and noted that the performance under encouragement was only slightly better than that under discouragement. Chase,[4] using 213 *Ss* of from 27–96 mo. of age, found that in the dynamometer test and a perforation test, it appeared highly probable that failure accompanied by reproof was a stronger motivator than success accompanied by praise. In one-third of the cases that Nelson noted,[5] the degree of success definitely corresponded to the amount of praise given on three successive trials of the Seguin Test.

Praise and reproof have also been observed to improve certain verbal activities. Gilchrist[6] found that his 'praised group' of college students raised their average scores

* Accepted for publication July 8, 1937. From the Psychological Laboratory of the University of Pennsylvania.

[1] R. E. Leaming, Tests and norms for vocational guidance at the 15-year-old performance level, *Psychol. Clin.*, 14, 1922, 193-217.

[2] For a general discussion of the literature on memory span, see A. B. Blankenship, Memory span: A review of the literature, *Psychol. Bull.*, 35, 1938, 1-25.

[3] G. S. Gates and L. Q. Rissland, The effect of encouragement and discouragement upon performance, *J. Educ. Psychol.*, 14, 1923, 21-26.

[4] L. Chase, Motivation of young children, *Univ. Iowa Stud.* Child Welfare, 5, 1932 (no. 3), 1-119.

[5] J. F. Nelson, A preliminary report on some uses of the psychological test situation for studying personality difference, in D. S. Thomas and associates, Some new techniques for studying social behavior, *Teach. Coll., Child Dev. Mong.*, No. 1, 1929, 1-213.

[6] E. P. Gilchrist, Extent to which praise and reproof affect a pupil's work, *School & Soc.*, 4, 1916, 872-874.

in the Courtis English Test about 80%, while the 'reproved group' showed no gain. Briggs found that 87% of his *Ss* (high school pupils) made better grades in course examinations after commendation and encouragement.[7] Even in such a reliable type of activity as an intelligence test, praise and reproof have been found to affect performance. Using the National Intelligence Examination on 408 grade school children, Hurlock concluded that praise and reproof were of approximately equal value as motivators.[8] In another study in which she made use of the same test,[9] she observed that incentives in the form of praise and reproof raised the average I.Q. scores 7 points, while less than 1 point improvement resulted from practice alone.

## PROBLEM

The problem of the present research was twofold: (1) to determine the effect of praise and reproof upon memory span performance; and (2) to determine the effect of praise and reproof upon memory span reliability.

## METHOD AND PROCEDURE

*Materials.* The materials used were sets of digits varying in length from 5 to 13 numbers. These digits were read aloud to each group of subjects at a rate slightly faster than 1 per sec. The usual precautions in presenting a series of digits for a test of memory span were followed; no emphasis was placed on any of the digits; no digits were placed in their natural order; there was an avoidance of rotation or of any numbers that suggested multiplication, division, addition or substraction; and no historical sequence of numbers was used.[10] Only forward memory span was tested; recall was in the same order as that in which the stimuli were presented. At each stimulus-value, two series of digits were presented. The three groups of *Ss* used were each tested twice, at an interval of one week, a different series of digits being used each week.

*Instructions.* All of the groups were given standard instructions for the first week: the second week, however, instructions for the three groups varied. The control group was given only the instructions received the previous week. The 'approved' group was given the following instructions before the second testing period:

> We have found that this group has done exceptionally well on the memory span test which we gave last week. We are so pleased, as well as amazed, by the excellent scores attained, that we would like you to try the test once more for us. As before, we request full coöperation.

The 'reproved' group received the following directions preceding their second week's test:

---

[7] T. H. Briggs, Praise and censure as incentives, *ibid.*, 26, 1927, 596-598.

[8] E. B. Hurlock, The value of praise and reproof as incentives for children, *Arch. Psychol.*, 11, 1924 (no. 71), 1-78.

[9] Hurlock, The effect of incentives upon the constancy of the I.Q., *Ped. Sem.*, 32, 1925, 422-434.

[10] For a discussion of precautions to be followed in a test of memory span, see A. Binet and T. Simon, Methodes nouvelles pour le diagnostic du niveau intellectuel des anormaux, *Année psychol.*, 12, 1905, 191-244; T. L. Bolton, The growth of memory in school children, this JOURNAL, 4, 1893, 362-380; B. Bourdon, Influence de l'age sur la memoire immediate, *Rev. Philos.*, 38, 1894, 148-167; R. A. Brotemarkle, Some memory span test problems, *Psychol. Clin.*, 15, 1924, 229-258; S. W. Calhoon, Influence of length of lists upon ability immediately to reproduce disconnected word series auditorially presented, *J. Exper. Psychol.*, 17, 1934, 723-738.

We have found that this group has done exceptionally poorly on the memory span test we gave last week. As a result, the figures secured are practically meaningless, and we are convinced that the poor scores were due primarily to lack of coöperation. Inasmuch as the experiment takes but a few moments, we think we are justified in requiring you to try the test again. Please coöperate with us this time.

*Subjects.* The *Ss* tested were all undergraduate women of Beaver College.[11] The control group, composed of 44 students, had an age-range of from 18–22 yr., the average being 20.45 yr. The approved group of 43 *Ss* had an age-range of from 17–23 yr., the average being 19.72 yr., while the reproved group, consisting also of 43 *Ss*, had an age-range of from 18–22 yr., the mean being 19.75 yr. To make sure that age would not be a factor in causing the groups to be non-comparable, a correlation was secured between age and memory span performance for 86 of the students. The coefficient of 0.08(.04)[12] indicates that there is no significant relation between age and memory span for the group studied,[13] and that the groups would not be non-comparable for this reason.

## Results

Table I shows the mean number of digits reproduced by each of the three groups for each test period. The table indicates that the control group showed the greatest

TABLE I

Showing Mean Number of Digits Reproduced on Each Test

| Group | Mean | |
|---|---|---|
| | Test 1 | Test 2 |
| Control | 7.86(1.41) | 8.41(1.64) |
| Approved | 7.86(1.92) | 8.12(1.60) |
| Reproved | 7.84(1.44) | 7.89(1.44) |

gain. Table II presents the numerical value of these gains, and shows the significance of the differences between the first and second tests for each of the three groups. In terms of the average number of digits gained, they are in all cases quite small, and the critical ratios indicate that in no case are the gains statistically significant. (The critical ratio should approach 3.00 to show a statistically significant difference.) In terms of percentage gain, the control group gained most, having an average gain of about 7%, while the approved group gained 3.3% and the reproved group

TABLE II

Showing the Significance of the Gain in Digits in Each Group

| Group | Av. gain | $\sigma$ gain | Critical ratio |
|---|---|---|---|
| Control | .55 | 2.16 | .25 |
| Approval | .26 | 2.49 | .10 |
| Reproval | .05 | 2.09 | .02 |

[11] The writers take this opportunity to express their sincere gratitude to Professor Elinor J. Barnes, who made arrangements for the experiment, and to the students who coöperated in it.

[12] Figures in parentheses following correlation coefficients or means refer to the sigma of the measure.

[13] This agrees with other experimental results. It has been found in general that memory span increases to the age of approximately 16 yr., and that it remains fairly constant thereafter. For a discussion of this question, see Blankenship, *op. cit.*, 14.

0.7%. Table III indicates, as might be expected from the preceding data, that the differences between the gains of the three groups are not significant, since in no case does the critical ratio approach 3.00. This lack of statistically significant differences betwen the groups may probably be accounted for by the fact that the *Ss* were college students, more sophisticated than the majority of the younger *Ss*

TABLE III

SHOWING THE SIGNIFICANCE OF THE DIFFERENCES OF GAIN OF EACH GROUP

| Groups compared | Diff. in gain | $\sigma$ Diff. | Critical ratio |
|---|---|---|---|
| Control-Approval | .29 | 2.17 | .13 |
| Control-Reproval | .50 | 2.22 | .23 |
| Approval-Reproval | .21 | 2.12 | .10 |

used by most researchers in the studies of motivation. While a gain of 7% for the control group may appear to be comparatively slight, it is not unusual, since some investigators have reported no apparent practice effects in the performance.[34]

Table IV presents the reliability coefficient for the performance of each group (the correlation of performance one and performance two). It will be observed that

TABLE IV

SHOWING RELIABILITY COEFFICIENTS FOR EACH OF THE GROUPS

| | |
|---|---|
| Control group | .47(.12) |
| Approved group | .88(.03) |
| Reproved group | .41(.13) |

the reliability coefficients for the control and reproved groups are quite similar, while that for the approved group is considerably higher. That the differences favoring the approved group are statistically significant is shown by Table V.

TABLE V

SHOWING SIGNIFICANCE OF THE DIFFERENCES INDICATED IN TABLE IV

| Groups compared | Diff. in rel. coeff. | $\sigma$ Diff. | Critical ratio |
|---|---|---|---|
| Control-Approved | .41 | .12 | 3.42 |
| Control-Reproved | .06 | .18 | .33 |
| Approved-Reproved | .47 | .13 | 3.62 |

Table V shows that the difference between the reliability coefficients of the control and reproved groups is not significant statistically, but that the difference between the control and approved groups or between the reproved and approved groups is statistically significant.

This difference favoring the approved group cannot be accounted for by the present investigators. Further studies along this line of research should yield interesting results. The reliability coefficient secured for the control group may be compared with reliability coefficients secured by other investigators who have experimented with the audito-vocal memory span for digits. Previous reliability coefficients secured range from 0.28[35] to 0.93.[36] If the difference in reliability

[34] Cf. for example, H. B. Reed, A repetition of Ebert and Meumann's practice experiment on memory, *J. Exper. Psychol.*, 2, 1917, 315-346.

[35] E. B. Bolton, The relation of memory to intelligence, *J. Exper. Psychol.*, 14, 1931, 37-67.

[36] C. Burt, Experimental tests of general intelligence, *Brit. J. Psychol.*, 3, 1909, 94-177.

coefficients favoring the approved group is substantiated, however, the implication for psychologists making clinical use of memory span is clear, for to secure reliable results in the test, all the *Ss* will have to be encouraged.

One might still question the possibility of a relationship between age and memory span gain. For this reason, correlation coefficients were secured between age and memory span gain in each of the three groups and these data are presented in Table VI. It can be seen immediately that there is only a chance relationship between

TABLE VI

SHOWING CORRELATION OF AGE AND MEMORY SPAN GAIN

| Group | Correlation coefficient |
|---|---|
| Control | .07(.15) |
| Approval | .11(.15) |
| Reproval | .12(.13) |

age and memory span gain in each of the groups, since all of the correlation coefficients secured in each group are so small and the sigma of each coefficient is so great that it is unnecessary to show by an additional table that the differences are statistically insignificant.

The composite results for the total group of *Ss* (numbering 130) on the first test of memory span were calculated in the forms of measures of central tendency. The mean memory span on test one for the total number of *Ss* was 7.85, while the median and mode were 7. These results are similar to those of Bingham,[17] who found that the median score of the Dartmouth men he tested was 7 digits. Humpstone,[18] however, secured a modal score of 8 digits for the college students he tested.

## CONCLUSIONS

As a result of the experimental investigation, the following conclusions were drawn.

*(1)* The effects of praise and reproof, as tested by the procedure used in this experiment, are negligible, so far as memory span is concerned. This is indicated by the fact that gains in memory span for each of the groups (control, approved, and reproved) were negligible, and differences between the gains of each group were not statistically reliable.

*(2)* In the groups tested, the approved group demonstrated a significantly more reliable memory span performance (0.88) than the reproved group (0.41) or the control group (0.47), while the difference in reliability coefficients between the last two groups was not statistically significant.

*(3)* The low correlation coefficients secured in each of the three groups (the coefficients ranged from 0.07 to 0.12) indicated that there was no relation between memory span gain and age in the groups tested. Preliminary work had shown that there was no relationship between age and memory span performance in the group tested (the coefficient was 0.08).

---

[17] W. V. Bingham, Some norms of Dartmouth men, *J. Educ. Psychol.*, 7, 1916, 129-142.

[18] H. J. Humpstone, Memory span tests, *Psychol. Clin.*, 12, 1919, 192-200.

# EFFECT OF STIMULUS-TEXTURE UPON APPARENT WARMTH AND AFFECTIVE VALUE OF COLORS

By MILES A. TINKER, University of Minnesota

In practically all experiments on affective value of color and on meanings of color such as warmth of colors, exciting *vs.* subduing colors, etc., the stimulus-objects have consisted of colored papers or colored lights. These colors have a surface texture which is markedly different from those found in most life situations. In nature, colored objects vary considerably in surface texture from that of colored papers. This is equally true for colored clothing and to a certain extent in paintings. One of few life situations where surface texture of colored objects approaches that of colored papers used in the laboratory is in magazine advertisements.

The purpose of this experiment is to investigate the effect of stimulus-texture upon apparent warmth and upon affective value of colors in two situations where the surface-texture is considerably different. The two types of stimulus-objects are, respectively, cloth and coated paper.

At the outset, it may be advisable to define clearly what is meant by apparent warmth of colors. There have been two important approaches employed in investigating the effect of color upon perception. In the first, an attempt has been made to discover whether color affects sensory perceptions such as cutaneous warmth of objects[1] and kinesthetic perception in lifting weights.[2] Results in these experiments have been highly inconsistent, unreliable and inconclusive. No consistent trend in effect of color upon sensory experience has been demonstrated. In the second type of approach the *apparent* effects of color have been studied by the "looks" method.[3] That is, the subject views the colored stimulus-objects and judges if one looks heavier or lighter than another, or larger or smaller than another, or warmer or cooler than another, etc. These experiments have yielded consistent and reliable trends of color effects. When judgments are made in terms of how a colored object looks, *apparent* effects are noted. These two types of experimental approach should be clearly differentiated. It is the latter approach that is employed in our study of apparent warmth.

*Materials.* There were two series of stimulus-objects in this experiment. The first consisted of 11 colors on Milton-Bradley coated paper: red, blue, yellow, green, orange, violet, blue-green, yellow-green, gray, white and black. In the second were comparable colors on cloth: red (claret), blue (royal), yellow (forsythia), green (emerald), orange (goya red), red-orange (romany red), violet (petunia), blue-

* Accepted for publication January 29, 1938. Thanks are due M. Diddy and M. Greenham who helped collect the data in this experiment.

[1] M. F. Morgensen and H. B. English, The apparent warmth of colors, this JOURNAL, 37, 1926, 427-428.

[2] C. D. Taylor, Visual perception versus visual plus kinaesthetic perception in judging colored weights, *J. Gen. Psychol.*, 4, 1930, 229-246; J. E. DeCamp, The influence of color on apparent weight, *J. Exper. Psychol.*, 2, 1917, 347-370.

[3] E. Bullough, On the apparent heaviness of colors, *Brit. J. Psychol.*, 2, 1907, 111-152; M. Monroe, The apparent weight of color and correlated phenomena, this JOURNAL, 36, 1925, 192-206; and references in footnotes 1 and 2 above.

green (amulet blue), yellow-green (chartreuse), gray (crane gray), black and white. The cloth was silk crepe manufactured by the Duplan Company. The names (claret, etc.) were the trade names given to the cloth samples. The cloth and paper stimuli were mounted between two pieces of gray cardboard, 8½ x 12 in. The middle of the top cardboard was cut away so that a panel of color, 4 x 7 in., was visible.

*Subjects.* The *Ss* in this experiment were 50 men and 50 women between the ages of 17 and 30 yr., mostly university students.

*Procedure.* On one day *S* was asked to place the cloth stimuli in rank order (1) according to apparent warmth (meaning or association of warmth) and (2) according to preference. On the second day the same was done for the paper stimuli. Cloth was employed first by one *S* and second by the next, etc. The sequence of rankings for apparent warmth and for affective value were also systematically rotated. Observation was in daylight illumination. All the *Ss* were asked to make their ranking in terms of the samples as seen, *i.e.* they were not to think of objects—a purple tie, a yellow scarf, etc.

The reliability of the ranking was determined by the split-half method. The mean ranks assigned by 25 men were compared to those given by the other 25, and likewise for the women. The rank-difference method of correlation yielded the following reliability coefficients:

| | |
|---|---|
| Cloth—Warmth: | Men, .94; Women, .95 |
| Paper—Warmth: | Men, .96; Women, .93 |
| Cloth—Preference: | Men, .92; Women, .96 |
| Paper—Preference: | Men, .93; Women, .83 |

These coefficients show that the rank orders were established with a high degree of reliability.

*Results.* The results are considered for the men alone, the women alone, and for the combined group of 100 *Ss*. The mean ranks for apparent warmth of colors are listed in Table I. The ranking of the cloth and of the paper are presented side by side for comparison. The successive mean ranks for each type of stimulus-object reveal in general a fair degree of separation. It is obvious that the *Ss* assigned different degrees of apparent warmth to the separate colors. The colors were considered to

TABLE I

MEAN RANKS ASSIGNED TO CLOTH AND PAPER STIMULI FOR WARMTH OF COLOR
(N= 100; 50 men and 50 women)

| Color | Men | | Women | | Both | |
|---|---|---|---|---|---|---|
| | cloth | paper | cloth | paper | cloth | paper |
| Red-orange | 2.3 | — | 2.2 | — | 1.6 | — |
| orange | 3.2 | 2.5 | 2.7 | 2.6 | 2.5 | 2.6 |
| red | 4.1 | 1.9 | 3.7 | 2.2 | 3.9 | 2.1 |
| violet | 5.1 | 6.2 | 5.1 | 5.1 | 5.1 | 5.7 |
| blue | 5.8 | 5.7 | 6.5 | 6.3 | 6.1 | 6.0 |
| green | 6.2 | 5.4 | 6.4 | 6.3 | 6.3 | 5.9 |
| blue-green | 7.3 | 7.3 | 7.9 | 7.5 | 7.7 | 7.4 |
| yellow | 7.4 | 4.9 | 6.2 | 4.8 | 6.8 | 4.8 |
| black | 7.6 | 8.2 | 8.2 | 7.1 | 7.9 | 8.0 |
| yellow-green | 8.4 | 6.1 | 8.5 | 6.4 | 8.4 | 6.2 |
| gray | 10.0 | 8.6 | 9.9 | 8.2 | 10.0 | 8.4 |
| white | 10.1 | 9.1 | 10.8 | 9.7 | 10.5 | 9.4 |

have more apparent warmth than the achromatic stimuli, since the black, gray and white papers and cloths tended consistently to be at the cool end of the series. With the colors the greatest apparent warmth was assigned to the reds and oranges, the greatest apparent coolness to blue-green and yellow-green. Blue and green occupied an intermediate position.

TABLE II

MEAN RANKS ASSIGNED TO CLOTH AND PAPER STIMULI FOR AFFECTIVE VALUE OF COLOR (N = 100; 50 men and 50 women)

| Color | Men | | Women | | Both | |
|---|---|---|---|---|---|---|
| | cloth | paper | cloth | paper | cloth | paper |
| Blue | 2.3 | 2.4 | 4.3 | 3.3 | 3.3 | 2.9 |
| green | 4.5 | 3.9 | 5.9 | 5.3 | 5.2 | 4.6 |
| violet | 4.6 | 6.5 | 6.9 | 5.5 | 5.8 | 6.0 |
| red-orange | 4.9 | — | 4.0 | — | 4.4 | — |
| red | 5.8 | 3.8 | 6.1 | 5.0 | 6.0 | 4.4 |
| orange | 6.0 | 4.9 | 3.9 | 4.3 | 4.9 | 4.6 |
| blue-green | 6.1 | 6.5 | 5.4 | 5.5 | 5.7 | 6.0 |
| white | 7.4 | 6.4 | 8.9 | 7.3 | 8.1 | 6.9 |
| black | 8.2 | 8.3 | 7.6 | 7.2 | 7.9 | 7.8 |
| gray | 8.7 | 8.4 | 9.1 | 9.0 | 8.9 | 8.7 |
| yellow | 9.1 | 6.6 | 6.6 | 5.4 | 7.8 | 5.9 |
| yellow-green | 9.3 | 8.1 | 9.3 | 7.6 | 9.3 | 7.9 |

In Table II are presented the data on affective value of the colors. The orders of preference are similar to those reported in other experiments.

The effect of stimulus-texture upon the rankings was obtained by correlating the ranks assigned to the cloth stimulus-objects with those assigned to the paper. The size of the correlations will show to what degree the ranks varied in the different comparisons. The coefficients are listed in Table III. In the same table are given the relationship between affective value and apparent warmth for cloth and paper. The correlations for the men are consistently lower than for the women in all comparisons. For men, women, and the combined group, the correlations between results from paper and cloth are 0.77, 0.93, and 0.90 respectively for apparent

TABLE III

INTERCORRELATIONS

| Comparison | Men | Women | Both |
|---|---|---|---|
| Warmth: Paper vs. Cloth | .77 | .93 | .90 |
| Preference: Paper vs. Cloth | .83 | .90 | .86 |
| Paper: Warmth vs. Preference | .56 | .74 | .72 |
| Cloth: Warmth vs. Preference | .68 | .75 | .79 |

warmth; and 0.83, 0.90, and 0.86 respectively for affective value. These results indicate that variations in surface texture of stimulus-objects have little effect upon apparent warmth and affective value of colors. The rankings are approximately the same for cloth as for paper. The women, however, are more consistent from one surface texture to another than are the men.

In the last two rows of Table III are given the correlations between rankings for apparent warmth and for affective value. The coefficients of 0.56 to 0.79 (Mean =

0.71) indicate a moderate degree of relationship. The warmer colors tend to be preferred. Again the coefficients for the men are slightly lower than for the women.

In most instances the sex differences in ranking are slight. When the rankings by men were correlated with those by women the following coefficients were obtained:

Apparent warmth for cloth stimuli, rho = .93
Apparent warmth for paper stimuli, rho = .95
Affective value for cloth stimuli, rho = .66
Affective value for paper stimuli, rho = .84

The only comparison which appears to show some tendency toward a sex difference is in ranking the cloth stimuli for affective value. This correlation of only .66 between color preference for men and for women indicates considerable discrepancy in rankings assigned by the two sexes.

*Conclusions.* The results of this experiment warrant the following conclusions: (1) Rankings for apparent warmth and for affective value of color can be established with a high degree of consistency. (2) Surface texture, as represented by coated paper versus cloth, has little or no effect upon apparent warmth or affective value of colors. (3) Warm colors tend to be preferred over cool colors or achromatic stimuli. (4) The ranking for apparent warmth of colors by men is practically identical to that for women. There is, however, a slight tendency toward sex differences in color preferences.

# THE PITCH OF A COMBINATION OF TONES

By PAUL R. FARNSWORTH, Stanford University

In 1923 the writer undertook a study of the phenomenal pitch of dyads. Before that date the issue had been clouded by two mutually exclusive sets of results—those of Stumpf[1] and those of Valentine.[2] Stumpf held that the pitch of the lower component of a dyad was more dominant than that of the higher component, while Valentine held that it was the pitch of the higher component which was the more dominant. In his earlier paper,[3] the writer had attempted to reconcile these conflicting sets of data by showing that while it was highly probable that Stumpf's subjects were musically naïve it was definitely known that Valentine's were musically trained. These latter, then, had been trained in melody hunting. This in itself should tend to emphasize the pitch of the upper component of each dyad, or so it seemed to the present writer. In the same paper, data were offered of a clinical sort which appeared to substantiate this theory that one's training may alter one's pitch reactions. The untrained subjects in the experiment reacted in the main as did Stumpf's; those trained in melody hunting as did Valentine's.

The writer's clinical data offered in 1923 were assembled from only 10 *Ss*. As they gave little information concerning the dyad pitch-discrimination of the average person (average in a musical sense) it was thought advisable to conduct similar experiments on a fairly large number of college students.

*Subjects.* The *Ss*, 276 in number, were drawn at random from classes in psychology. Half (138) were arbitrarily classified as musical, *i.e.* they had had from 1 to 12½ years of musical training of some kind (median 2½ yr.), and half as unmusical, *i.e.* they had had no musical training or less than a year of training. The *Ss*, musical and unmusical, served together in groups of 10 to 20.

*Directions.* The following directions were given the *Ss*.

"In four experiments you will hear a single tone followed by two dyads (two tones played simultaneously). You are to check the dyad which more closely resembles the single tone in pitch. There is no correct answer. We merely wish to get your impressions in this matter."

*Procedure.* Every one of the four experiments consisted of a single tone and two dyads. The single tone was sounded first, held for 1 sec., then 2 sec. later the first dyad was sounded, held in turn for 1 sec., and then 2 sec. later the second dyad was sounded and held for 1 sec. An interval of 5 sec. lapsed between successive experiments.

All the tones sounded were c's of different octaves, and the dyads contained the single tone as the upper or lower member. The first and third and the second and fourth experiments were alike respectively except for the fact that the dyads were presented in them in reverse order. This precaution was taken to eliminate a pos-

* Accepted for publication August 16, 1937.

[1] Carl Stumpf, *Tonpsychologie*, 1890, 384.

[2] C. W. Valentine, The aesthetic appreciation of musical intervals among school children and adults, *Brit. J. Psychol.*, 6, 1913, 190-216.

[3] P. R. Farnsworth, Notes on the pitch of a combination of tones, *ibid.*, 15, 1924, 82-85.

sible time-error. The tones were given automatically at 'dance' intensity by a Duo-Art Pianola.

In Experiment 1, the single tone was 512 ~, the first dyad was 256 ~ and 512 ~, and the second dyad 512 ~ and 1024 ~. In Experiment 2, the single tone was 1024 ~, the first dyad 1024 ~ and 2048 ~, and the second dyad 512 ~ and 1024 ~. In Experiment 3, the single tone was again 512 ~ but the first dyad was 512 ~ and 1024 ~, and the second was 256 ~ and 512 ~. In Experiment 4, the single tone was 1024 ~, the first dyad 512 ~ and 1024 ~, and the second 1024 ~ and 2048 ~. If the *Ss* chose the higher dyad they were designating the bass element as that element was of the pitch of the single tone; if they chose the lower dyad they were designating the soprano element as in that instance the soprano element was of the pitch of the single tone.

When those experiments were completed, the *Ss* were given the following instructions for Experiment 5.

"Next you will hear two dyads. Note on your paper which of the two seems to have the higher pitch."

The sounds compared consisted of the dyad 512 ~ and 1024 ~, held for 1 sec., followed 2 sec. later by the dyad 645 ~ and 767 ~, also held for 1 sec. In terms of the musical scale, the first dyad is an octave of c's, and the second is an interval of the third, e-g. It should be noted that the boundaries of the second dyad—in terms of intervals—are within, and equidistant from, those of the first. The designation of the larger of the two dyads as the one having the higher pitch will be a 'vote' for the soprano element; conversely, a ballot for the smaller of the dyads will serve as a choice of the bass element.

*Results.* The results, computed separately for the musical and unmusical *Ss*, are given in Table I which shows the percentage of selections favoring the soprano element of the dyad. The results from Experiments 1 and 3 and from 2 and 4 respectively are grouped together as they are identical except for time-order.

TABLE I

PERCENTAGE OF REPORTS (FROM ALL *Ss*) FAVORING THE SOPRANO ELEMENT

(138 *Ss* in each group)

| Exper. | Group | % | Diff. | S.E. | Crit. Ratio |
|---|---|---|---|---|---|
| 1 & 3 | Music. | 47.8 | 4.0 | 5.9 | 0.68 |
| | Unmus. | 43.8 | | | |
| 2 & 4 | Music. | 36.2 | 5.0 | 5.6 | 0.89 |
| | Unmus. | 31.2 | | | |
| 5 | Music. | 54.4 | 24.0 | 5.7 | 4.21 |
| | Unmus. | 30.4 | | | |

With the exception of Experiment 5 the critical ratios are extremely small; yet the differences are all in the expected direction (fewer of the unmusical *Ss* favored the soprano element). In Experiments 1 and 3 only a bare majority in both groups favored the bass element (Stumpf's theory). In Experiments 2 and 4 the percentages are greater. In Experiment 5 it can be seen that while a bare majority among the musical *Ss* favored the soprano element, considerably more among the unmusical favored the bass.

Direct questioning of the *Ss* and their spontaneous remarks revealed quite

clearly that the instructions of Experiment 5 (Which dyad is higher in pitch?) seemed easier to follow than those of the other experiments (Which dyad more closely resembles the single tone in pitch?). The former seemed a more meaningful question. Many were bothered when they noticed that each dyad in the earlier measures contained the single tone which preceded it.

The division of the *Ss* into musical and unmusical was, it will be recalled, made upon an arbitrary basis. To contrast still more the effects of training, the tails of the distribution were compared. The reports of the *Ss* who had had 4 years or more of music training are compared in Table II with those with no musical training. It should be recognized that even this division was artificial as a number of the musical group had probably profited little from their work in music while most if not all of the unmusical *Ss* had quite obviously had some practice in melody hunting. One can only guess as to what the differences between the reports would have been had truly musical and unmusical *Ss* been employed.

TABLE II

PERCENTAGE OF REPORTS (FROM MOST AND LEAST MUSICAL *Ss*) FAVORING THE SOPRANO ELEMENT

| Exper. | Group | No. Ss | % | Diff. | S.E. | Crit. Ratio |
|---|---|---|---|---|---|---|
| 1 & 3 | Music. | 47 | 48.9 | 5.5 | 8.8 | 0.63 |
| | Unmus. | 98 | 43.4 | | | |
| 2 & 4 | Music. | 47 | 37.3 | 6.2 | 8.4 | 0.74 |
| | Unmus. | 98 | 31.1 | | | |
| 5 | Music. | 47 | 57.5 | 34.0 | 8.3 | 4.10 |
| | Unmus. | 98 | 23.5 | | | |

The data of Table II resemble very closely those of Table I. While the differences are slightly larger, the decreases in population make the critical ratios even smaller than those shown in the first table.

In an attempt to check the theory that sopranos with their greater experience in melody hunting should favor the soprano element more than other women-singers, and that the latter with somewhat less experience should still favor the soprano element more than men-singers, the five experiments described above were presented to 126 trained vocalists. Two members of the Stanford University music faculty,[4] and the present writer gave the experiments to these *Ss* in small groups. Among the *Ss* there were 33 sopranos, 38 other women and 55 men. All the *Ss* were members of glee clubs or choirs. Their singing experiences varied greatly. Because it was impossible to get these people to the room in which the Duo-Art Pianola was kept, the tones were given on a piano by hand and approximated the same order and time sequences as in the earlier experiments. As a number of differences in procedure probably occurred, it will be unwise to compare directly the percentages of Table III with those of Tables I and II.

---

[4] The writer is grateful to Professors Warren Allen and Raymond Kendall for their aid in this study.

TABLE III

PERCENTAGE OF REPORTS (FROM TRAINED VOCALISTS) FAVORING THE SOPRANO ELEMENT

| Exper. | Group | No. Ss | % | Groups | Diff. | S.E. | Crit. Ratio |
|---|---|---|---|---|---|---|---|
| 1 & 3 | Sopr. | 33 | 56.1 | Sopr. & oth. women | 10.0 | 11.8 | 0.85 |
| | Oth. wom. | 38 | 46.1 | Sopr. & men | 14.8 | 10.9 | 1.36 |
| | Men | 55 | 41.3 | Oth. women & men | 4.8 | 10.4 | 0.46 |
| 2 & 4 | Sopr. | 33 | 46.9 | Sopr. & oth. women | 6.1 | 11.7 | 0.52 |
| | Oth. wom. | 38 | 40.8 | Sopr. & men | 6.5 | 10.9 | 0.60 |
| | Men | 55 | 40.4 | Oth. women & men | 0.4 | 10.2 | 0.04 |
| 5 | Sopr. | 33 | 69.7 | Sopr. & oth. women | 14.5 | 11.3 | 1.29 |
| | Oth. wom. | 38 | 55.2 | Sopr. & men | 25.5 | 10.3 | 2.47 |
| | Men | 55 | 44.2 | Oth. women & men | 11.0 | 10.4 | 1.06 |

None of the critical ratios reaches the statistician's standard value of 3.00. Yet the differences are all in the expected direction, and are of considerable magnitude in Experiment 5. It is of interest to note that in only three instances are the percentages favorable to the soprano element and in two of the three cases the lead is only by a bare majority.

## SUMMARY

In these experiments the writer attempted to gather further data to prove or disprove the theory that training in melody hunting can influence pitch perception in dyads. Two procedures were employed. In the first the *Ss* were presented with a single tone followed by two octave dyads which contained this tone. They were asked which dyad more resembled the single tone in pitch. In the second procedure (which the *Ss* markedly preferred) two dyads were presented along with the request that the *Ss* decide which gave the higher pitch. The boundaries of the smaller of the dyads were within and equidistant from those of the larger.

The *Ss*, 276 in number, were subdivided into the relatively more and the relatively less musical. When the second procedure was employed it was apparent that the more musical *Ss* with their greater experience in melody hunting violated Stumpf's rule (that the pitch of the lower component of a dyad is more dominant than that of its higher element) even more than did the less musical students. When the first procedure was employed the two groups were found to differ but to a smaller and statistically insignificant degree.

In following through the notion that sopranos have had, on the average, more experience in melody hunting than other women singers, and that these latter have probably had more of such experience than have men singers, a number of vocalists were tested in the manner of the other experiments. In every comparison those who presumably had had greater experience in melody hunting violated Stumpf's rule more than did those with lesser experience. In practically all of the situations, however, the differences were not large. Yet the fact that they consistently appeared and the added finding that in some instances the soprano element was favored would seem to furnish further proof that Stumpf's early formulation must be modified to include the effects of training.

## MINOR STUDIES FROM THE PSYCHOLOGICAL LABORATORY OF CORNELL UNIVERSITY

### LXXXVI. The Effect of Light-Flashes During the Course of Dark Adaptation

By Laura K. Allen and Karl M. Dallenbach

Conflicting statements appear in the literature regarding the effect of flashes of light upon the subsequent course of dark adaptation. Aubert, for instance, states as early as 1865 that "the progress of dark adaptation was not only not impeded by flashes of light, but that, in fact, it was aided in some measure."[1] This observation, Nagel writes, "has since been confirmed many times. . . . That the instantaneous light of a match reflected from dark walls into the observer's eye lowers the threshold of the light sense considerably, as much perhaps as one-third, can easily be verified."[2] In opposition Parsons says that "a very short exposure to bright light, *e.g.* striking a match, causes a temporary fall without materially altering the course of the curve."[3]

Since these authors do not give the precise conditions under which their observations were made, it may very well be that the conflicting statements merely reflect a difference in the conditions of observation. Logically there are three possible answers to the question what is the effect of flashes of light upon the subsequent course of dark adaptation. The course may be (a) retarded, (b) hastened, or (c) unaffected. Any one of these results may be obtained if the effect of the flashes varies (1) with their intensity, (2) with their duration, and (3) with the point in the course of adaptation at which they are introduced. The effect may also be different (4) in monocular and binocular vision.

Every one of these variables must be considered separately and in various combinations before a final answer to our question may be returned. In the present study, however, we have limited ourselves to the consideration of (3). Observing monocularly and keeping the intensity and duration of the light-flashes constant, we have endeavored to determine the effect of varying the point in the adaptation-curve at which the flash is introduced.

#### Method and Procedure

*Apparatus.* The course of dark adaptation and the effect of flashes of light introduced at different-times was observed by means of a Nagel adaptometer[4] set up

* Accepted for publication February 1, 1937.

[1] H. Aubert, *Physiologie der Netzhaut*, 1865. Quoted from W. Nagel, Adaptation, twilight vision, and the duplicity theory, in Helmholtz's *Physiological Optics* (Eng. trans.), 2, 1924, 323.

[2] W. A. Nagel, *op. cit.*, 323.

[3] J. H. Parsons, *An Introduction to the Study of Colour Vision*, 2nd. ed., 1924, 53.

[4] For a description of this apparatus see W. A. Nagel, Zwei Apparate für die augenärtzliche Funktionsprüfung: Adaptometer und kleines Spektralphotometer (Anomaloskop), *Zsch. f. Augenhk.*, 3, 1907, 116-118. The apparatus is also described by W. Lehmann, Über Helladaptation, *Zsch. f. Sinnesphysiol.*, 41, 1907, 304 ff., and a second time by Nagel in his article in Helmholtz's *Physiological Optics*, op. cit., 315-317.

in a dark and semi-soundproof room. The exposure-field of this apparatus, a milk-glass plate, can be varied intensively from 1 to 1/80,000,000 parts and accurately evaluated from 1 to 1/8,000,000 parts. The stimulus-field used in this study was the area of an equilateral triangle,[5] 7.5 cm. in altitude, with base up and apex pointing directly downward.[6] The fixation-point, a weak red light 1 mm. in diam.,[7] was placed in the plane of the stimulus-field and 43 mm. above its upper edge.[8] Seated and with his head held comfortably and firmly in position by means of a wax biting board, *O* observed with his better eye. The biting board was adjusted for every *O* separately so that the fixation-point was directly in his monocular line of regard and the fixation-distance was exactly 1 m. The handle of the adaptometer, by turning which the intensity of the stimulus-light was varied, was extended so that *O* could make the adjustments from his chair.

The exposure-field was illuminated from within the apparatus by three 40-w. bulbs on a 110-volt D.C. line. The strength of current was dropped, however, and maintained throughout the experiment at 100 volts by means of a rheostat and voltmeter that were placed in a light box on *E*'s desk.

A large white screen (120 cm. broad and 102 cm. high), with apertures for exposure-field and fixation-point, was placed 1 m. in front of *O* in the plane of the apparatus. The screen not only concealed the adaptometer and its accessory apparatus, but also served as a preëxperimental adaptation-field and as a reflector of the light-flashes that were introduced during the course of the experiment.

Since the several *O*s, and the same *O* at different times, came to the experiments under very different states of light-adaptation—depending upon the brightness or dullness of the day and upon the previous use of their eyes—, every *O* was brought to a relatively uniform and constant state of adaptation by being required to fixate a black spot (1 mm. in diam.) at the center of the screen for 5 min. under the illumination of a 200-w. hydrogen-filled mazda lamp. This lamp, which was controlled from *E*'s desk, was in a reflector placed a little above and back of *O* so that neither the light itself, nor the shadows cast by it, could be seen by him in peripheral vision. The intensity of the light at the fixation-point was 48 foot-candles.

This light and screen were also used in producing the light-flashes that were introduced at various points in the course of adaptation. The duration of the flashes, controlled by a condenser-charge timer,[9] was 40 m.sec., and the points in the temporal course at which they were introduced were 5, 25, and 60 min.[10]

---

[5] We used a triangle because it is more easily distinguished from the idioretinal lights than either a circular or square field.

[6] This position brought the greatest part of the stimulus-field into the area of clearest vision.

[7] We used red because, as Nagel remarks, "it is more easily seen in the fovea centralis than in the peripheral parts of the retina, and hence the gaze is attracted by it" (*op. cit.*, 317).

[8] As the visual angle between the fixation-point and the upper edge of the stimulus-field was 8° 45′, the observations in this study were parafoveal.

[9] W. D. Moeder and J. G. Jenkins, A condenser-charge timer, this JOURNAL, 45, 1933, 742-744.

[10] We used an intense flash—the light of a 200-w. hydrogen-filled mazda lamp reflected from a white screen 1 m. in front of O—in this study on the principle that if a weak flash were to produce an effect, a strong flash would produce a greater effect. The principle may be wrong, but if so we would discover that in a subsequent study in which the intensity of the flash was alone varied.

*Instructions.* Immediately after the 5-min. period of light-adaptation the light was extinguished and the following instructions were read to *O*.

Place the blindfold over your poorer eye [this had been determined previously] and your mouth in the biting board [already adjusted] and fixate the red-point of light that appears directly in front of you. At the 'Ready, Now' signal, turn the handle of the apparatus (placed in a holder directly above your right hand) slowly to the left until you see a light below the fixation-point. You will be permitted to make small adjustments until you reach the point at which a light can just be seen. When this point has been reached, say "There," close your eye immediately and hold it closed until the next signal is given. Then repeat the procedure. A light may flash upon the screen at certain intervals; do not permit this to interfere with the course of the experiment but continue as before.

*Observers.* The *O*'s were Miss Martha L. Lemmon (*Le*), Mr. Edward A. Suchman (*Su*), Mr. Henry Werner (*Wr*)—graduate students in psychology—and the senior author (*Da*). All possessed normal vision and none wore glasses except *Da*, who dispensed with them during the observations as his vision was normal for distances beyond 16 in. *Da* and *Wr* had blue eyes, *Le* dark gray, and *Su* brown. *Da* and *Wr* were partially color blind. *Da* alone knew the purpose of the experiment.

*Procedure.* A modification of the method of limits, continuous presentation and ascending series alone, was employed in this study.[11] Descending series were omitted as they would have required superliminal stimulus-values, the intrusive effects of which we were investigating in this study. *E* set the intensity of illumination below *O*'s threshold and then, at the signal, *O* increased the intensity until he could just perceive the stimulus-light,[12] when he reported, and closed his eyes until the starting signal for the next series was given. During the interim, *E* recorded *O*'s setting, checked the voltage of exposure-lights, set the apparatus for the next series and then gave the signal, when the procedure was repeated.

*O* worked rapidly and ran as many series as possible during the course of the experimental hour. The number of series during the successive 5-min. periods of the experiment varied from 7–10 for *Da*, who was the slowest in making his adjustments, to 8–15 for *Wr*, who was the most rapid. The light-flashes, when given, were introduced without warning after the 'Ready, Now' signal and after *O* had opened his eye and had started to make the adjustments.

A long practice period, in which curves for the normal course of dark adaptation were obtained, was given every *O*. When the normal curves had become relatively uniform, indicating *O*'s stability of judgment and action, the following series of experiments, upon which the results of this study are based, were conducted. For every *O* except *Da*, 5 adaptation-curves of 40 min. each were obtained under every one of the following conditions: Series A, a flash after 25 min. of dark adaptation; Series B, a flash after 5 min. and a second flash after 25 min.; and Series C, normal, with no flash. For *Da*, 2 curves were obtained in Series B and C each, and 2 from experiments that were extended to 75 min. in one of which the light-flash was given after 60 min. These last curves with *Da* were secured because of Nagel's report that the effect is "particularly distinct when the observer has reached a practically stationary

[11] The same method was used by Nagel, *op. cit.*, 317.

[12] Our intensity-units convert into millilamberts when multiplied by the constant $10^{-7}$.

condition of adaptation after having been in the dark for an hour.[13] The different series were varied in haphazard order so that the *O*'s would not know which to expect.

## Results

*Normal curves.* The normal curves of dark adaptation (Series C) served as the basis of reference and comparison for determining the effect of light-flashes on the curves obtained in Series A and B. These normal curves, as the examples given in Fig. 1 show, are typical adaptation-curves, agreeing closely with those obtained by other experimenters.[14] The intensity required to yield a just noticeable light fell

TABLE I

Averages of the Just Noticeable Intensities for the Successive 5-Min. Intervals

| *O* | Series* | 5-min. intervals | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| *Le* | Normal | 725,871 | 32,406 | 14,343 | 7,789 | 6,068 | 5,237 | 5,187 | 5,184 |
| | A | 584,871 | 30,877 | 15,972 | 11,517 | 8,842 | 6,988 | 6,250 | 6,566 |
| | B | 669,016 | 38,130 | 18,262 | 9,559 | 7,036 | 7,756 | 5,596 | 4,751 |
| *Su* | Normal | 529,381 | 26,173 | 7,379 | 4,204 | 3,450 | 2,965 | 2,823 | 2,525 |
| | A | 554,235 | 38,588 | 10,230 | 4,959 | 3,728 | 3,406 | 2,760 | 2,444 |
| | B | 705,708 | 39,288 | 10,395 | 5,057 | 4,002 | 3,782 | 2,829 | 2,522 |
| *Wr* | Normal | 135,092 | 9,354 | 1,796 | 1,317 | 1,262 | 859 | 855 | 801 |
| | A | 313,127 | 12,294 | 1,867 | 1,260 | 1,057 | 1,105 | 838 | 794 |
| | B | 342,196 | 11,295 | 2,178 | 1,423 | 1,216 | 1,104 | 818 | 905 |

* In Series A the flash occurred after 25 min. of adaptation. In Series B one flash occurred after 5 min. and a second flash after 25 min.

at first rapidly and then more and more gradually until a level was reached beyond which the intensity declined very slowly if, indeed, it declined at all. The *O*s differed considerably, however, in the time required to reach this stationary level, in the intensity of that level, and in the variability of their adjustments during the same and at different experimental hours. *Da* and *Wr*, for example, attained the stationary level in from 10–12 min., whereas *Le* and *Su* required from 15–20 min. The final level attained by *Da* (1,200 intensity units) and *Wr* (600) was, moreover, much lower than that reached by *Le* (5,000) and *Su* (2,400).[15] Variability of adjustment was greatest for *Le*, smallest for *Da* and *Wr*, and intermediate for *Su*.

---

[13] *Op. cit.*, 323.

[14] *Cf.* S. Hecht, The nature of the photoreceptor: V. Dark-adaptation, *Foundations of Experimental Psychology*, 1929, 230 f., and the citations given there.

[15] It is worthy of note that *Da* and *Wr*, whose eyes are blue, attained considerably lower limits of dark adaptation than did *Le* and *Su*, whose eyes are dark gray and brown respectively. These results stand in opposition to those reported by H. Helson and J. P. Guilford (The relation of visual sensitivity to the amount of retinal pigmentation, *J. Gen. Psychol.*, 9, 1933, 58-76), who found that "it requires nearly twice as much light to attain the minimum visible for the 'average' blue-eyed subject as for the 'average' dark-eyed subject" (p. 65). Both of our blue-eyed *O*s, it should be recalled, were partially color-blind. It may be that this defect in their vision cuts across Helson and Guilford's generalization.

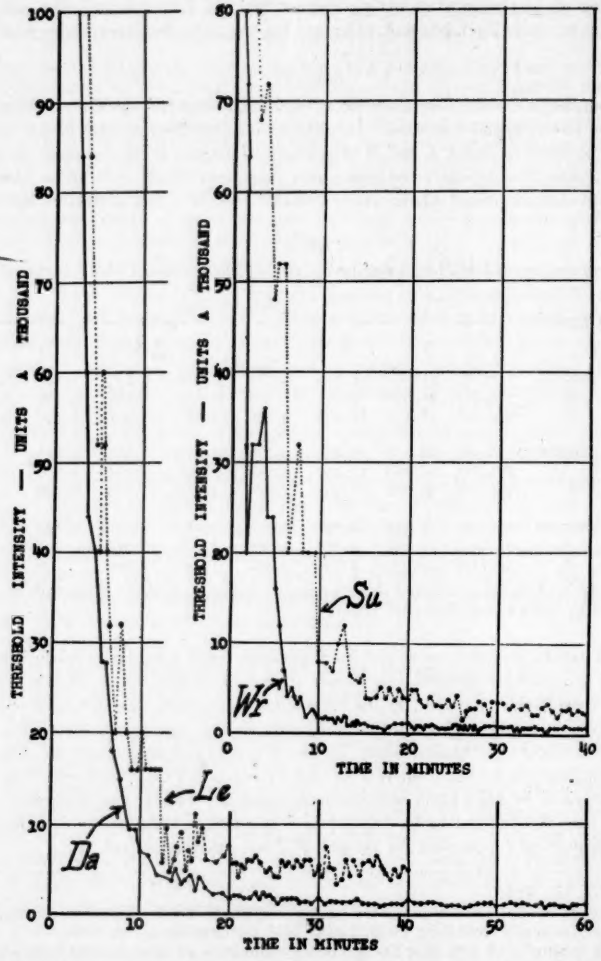


FIG. 1. SAMPLE NORMAL CURVES

*Flash curves.* Except for the momentary rises immediately following the introduction of the flashes of light, the flash-curves are very similar to the normal. The stationary level of adaptation is reached by the several *O*s in approximately the same time as in the normal series. The intensity at that level is also approximately the

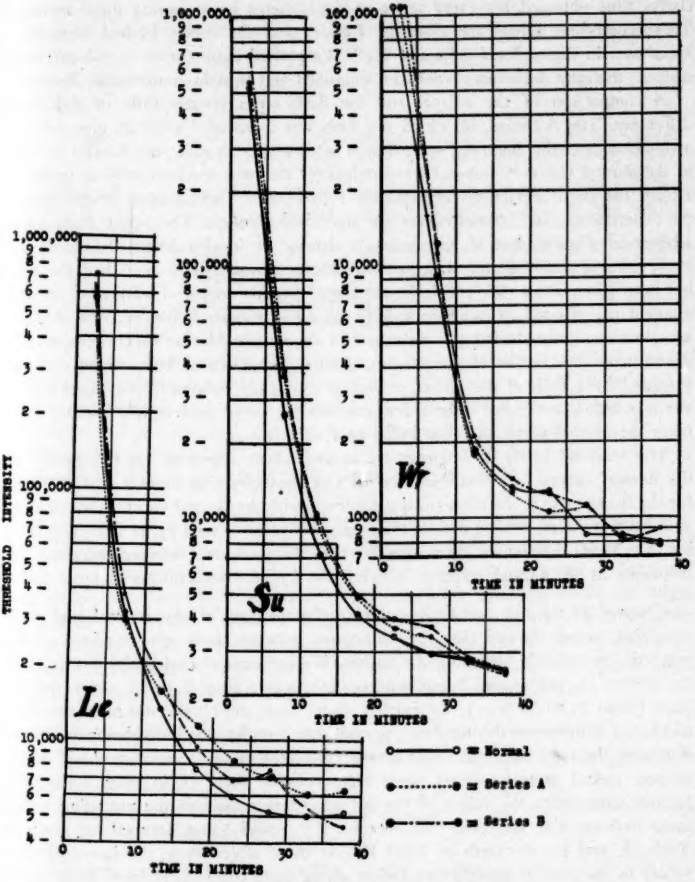


FIG. 2. COMPOSITE ADAPTATION CURVES

same, and the range of variability is of the same order, as in the normal series. The differences between the 'flash' and the 'normal' curves are no greater than the differences in the 'flash' curves or in the 'normal' curves themselves. There is nothing, therefore, in these results which indicates that flashes of the intensity and duration we used either hasten or retard adaptation.

*Smoothed curves.* In order to render comparison easy, we averaged the just noticeable intensity-values of every 5-min. interval, and then combined the 5 smoothed

curves thus obtained for every series of experiments by averaging these averages. These composite values are given in Table I and the curves plotted from them are shown in Fig. 2 for *Le*, *Su* and *Wr*.[36] We plotted these curves on a logarithmic scale so that any differences would be magnified and thus be more easily discerned.

A comparison of the normal and the flash series reveals little of significant difference. The A-curves, in which the flash was introduced after 25 min. of dark adaptation, and the B-curves, with flashes after 5 and 25 min., are similar in form to the normal C-curves and reach approximately the same terminal level of intensity. Except for slight deviations immediately following the introduction of the flashes, the experimental and normal curves are practically identical. The 5-min. flash delays adaptation slightly, but if adaptation is delayed it is also hastened because the curve after a few minutes is apparently where it would have been had the flash not been given at all. We conclude, therefore, that the course of adaptation, though momentarily affected, is not permanently altered by intense flashes of light of 40-m. sec. duration given during the early part of adaptation. Similar conclusions may be drawn from the results of the 25-min. flashes. The 25-min. flash, whether or not preceded by a flash at the 5-min. period, is uniformly followed by a slight rise in the adaptation-curve; but after a few minutes the curve falls rapidly until it overtakes the normal curve and then keeps pace with it.

The terminal levels of intensity are approximately the same for the 'flash' and the 'normal' curves. *Le's* final values varied 1,815 units from an average of 4,751 units for the B-curves to 6,566 units for the A-curves, with the normal C-curves terminating at a value between these extremes. *Su's* terminal values varied by 81 units and *Wr's* by 111. These differences are no greater than the variations between successive adjustments in the normal series or between the terminal adjustments in those series.

*Comparison of the post-flash intervals.* Comparison of the normal and flash smoothed curves showed that the differences between them were confined to the intervals immediately following the flashes. We consequently selected those parts of the curves, *i.e.* the second 5-min. interval (extending from 5 to 10 min.) and the sixth (from 25 to 30 min.), for further study. Since the *Os* did not make the same number of adjustments during these intervals, nor complete the individual adjustments at exactly the same time, the crude data were treated by the Vincent method[37] and 6 equally spaced intensity-values were computed for every *O* in every series. For further comparison, the values of the last adjustments made in the preceding 5-min. intervals were also recorded. The data for the second 5-min. interval are given in Table II, and for the sixth in Table III. As these tables show, the just noticeable values in the normal experiments follow along quite consistently from those in the preceding interval, but in every instance the value in the flash series shows abrupt increases, *e.g.* after the 5-min. flash from 88,800 to 108,000 units and after the 25-min. flash from 3,320 to 6,060 (*Su*, Tables II and III). The increase in every instance was of brief duration. The just noticeable values are comparable to the normal, in

[36] The results of *Da*, while corroborating those of the other *Os*, are not reported here as they were obtained under slightly different conditions.

[37] For a description of this method see S. B. Vincent, The function of the vibrissae in the behavior of the white rat, *Behav. Monog.*, 1, 1912 (no. 5), 16 f. *Cf.* also C. L. Kjerstad, The form of the learning curves for memory, *Psychol. Monog.*, 26, 1919, (no. 116), 26.

TABLE II

JUST NOTICEABLE INTENSITIES DURING THE SECOND 5-MIN. INTERVAL OF THE EXPERIMENTS

| O | Series | Av. last measurements at 5 min. | Vincent values | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | 1 | 2 | 3 | 4 | 5 | 6 |
| *Le* | Normal | 65,600 | 56,000 | 42,400 | 25,100 | 26,200 | 24,600 | 24,000 |
| | B | 72,800 | 79,800 | 41,200 | 37,000 | 31,700 | 30,800 | 26,400 |
| *Su* | Normal | 59,200 | 46,000 | 41,640 | 23,980 | 18,360 | 16,800 | 12,800 |
| | B | 88,800 | 108,000 | 45,900 | 34,900 | 30,000 | 24,000 | 18,400 |
| *Wr* | Normal | 20,700 | 14,000 | 9,540 | 5,540 | 4,140 | 3,140 | 2,000 |
| | B | 25,600 | 30,400 | 16,300 | 10,200 | 4,720 | 3,200 | 2,640 |
| *Da* | Normal | 32,000 | 26,000 | 22,100 | 16,500 | 14,100 | 10,000 | 6,400 |
| | B | 24,000 | 26,000 | 12,000 | 14,600 | 12,000 | 10,300 | 6,500 |

many instances, at the second determination and, in every instance, before the end of the 5-min. interval is reached. If adaptation is retarded as a result of the flash it is also hastened, because within a short time—from 2 to 4 min.—it has regained its lost ground and is equal again to the adaptation in the normal series. In view of the temporary character of the effects, we conclude with Parsons that the course of adaptation is itself not altered by flashes of bright light.

*After-sensations.* All the *O*s reported that the flashes, at both the 5- and 25-min. intervals, were followed by intense after-sensations which were positive at first and which then fluctuated with rapidly decreasing intensity between the positive and negative phases. These after-experiences lasted for a few minutes—about the length of time that the adaptation-curves deviated from the normal. When their intensity was greatest, the flash-curves deviated most from the normal. As their intensity decreased, the flash-curves approached more and more closely to the normal until, with the disappearance of the after-sensations, they approximated it.

TABLE III

JUST NOTICEABLE INTENSITIES DURING THE SIXTH 5-MIN. INTERVAL OF THE EXPERIMENTS

| O | Series | Av. last measurements at 25 min. | Vincent values | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | 1 | 2 | 3 | 4 | 5 | 6 |
| *Le* | Normal | 5,440 | 4,080 | 5,520 | 6,560 | 5,080 | 5,560 | 6,040 |
| | A | 5,220 | 14,500 | 7,620 | 6,240 | 6,220 | 5,780 | 6,440 |
| | B | 7,620 | 14,600 | 6,820 | 6,980 | 5,600 | 6,400 | 7,600 |
| *Su* | Normal | 3,120 | 3,480 | 3,280 | 2,720 | 2,980 | 2,640 | 2,840 |
| | A | 3,480 | 6,060 | 3,220 | 3,380 | 2,640 | 2,780 | 3,080 |
| | B | 3,320 | 6,060 | 3,380 | 2,640 | 2,780 | 3,080 | 3,782 |
| *Wr* | Normal | 1,060 | 840 | 870 | 680 | 920 | 870 | 880 |
| | A | 840 | 2,360 | 1,090 | 990 | 840 | 1,040 | 1,055 |
| | B | 1,080 | 2,240 | 1,140 | 1,180 | 970 | 910 | 1,080 |
| *Da* | Normal | 1,800 | 1,900 | 1,500 | 1,790 | 1,430 | 1,230 | 1,400 |
| | B | 2,100 | 2,700 | 1,350 | 1,150 | 1,000 | 1,000 | 1,200 |

*Supplementary results.* The supplementary experiments with *Da,* in which adaptation was continued for 75 min. and the light-flash was given after 60 min., yielded nothing new. The flash was followed, as in the experiments with the shorter intervals, by a rise and fall in the adaptation-curve, but within a few minutes the curve was again normal. The effects at this interval, the most favorable according to Nagel, did not differ in kind or in degree from those produced at the shorter intervals.

## SUMMARY

The results of this study, in which we sought to determine the effect of bright flashes of light upon the subsequent course of dark adaptation, may be briefly summarized.

*(1)* Flashes of light, under the conditions of our study, do not alter the course of dark adaptation.

*(2)* This conclusion holds equally for flashes that are introduced after short (5-min.) periods of adaptation when the curve is falling rapidly, after periods of intermediate length (25-min.) when the curve is flattening, or after long periods (60-min.) when the curve has reached a stationary level.

*(3)* The flashes have momentary effects which are similar at all parts of the adaptation-curve. After a flash the curve rises suddenly and briefly, recovers rapidly and returns quickly to the level at which it would have been had the flash not occurred.

*(4)* An earlier flash does not influence the effect of a later flash, providing the interval between them is sufficiently long for adaptation to recover.

## APPARATUS

# THE PYRAMID PUZZLE: A USEFUL DEVICE IN STUDYING THOUGHT

By ROBERT H. SEASHORE, Northwestern University

The experimental psychology of thinking has long been retarded by the lack of a suitable type of problem for analysis and measurement. Very few experiments appear in either laboratory manuals or research articles. The present paper outlines a suitable problem together with preliminary results and suggestions for a more systematic development and application.

The pyramid puzzle or disk transfer problem, supposedly coming from the early Egyptians, had its first scientific application as an illustrative problem in mathematics. Previous psychological applications include Peterson's comparison of the learning abilities of white and negro children,[1] Seashore and Seashore's demonstrations of insight and trial and error,[2] and Cook's study of the effects of complication in learning.[3] In none of these studies, however, have the full possibilities of the problem been realized.

For convenience of other experimenters a rather full outline of procedure is given in form suitable for either individual or group experiment. For most purposes it is desirable to give a preliminary definition of terms involved in problem solving so that the observers may be able to report their findings efficiently. This outline can be greatly shortened or expanded depending upon the amount of detail desired.

Almost any study of learning involves at least two major variables: (1) *the nature of the trials* by which an individual solves a problem, whether random or purposive; and (2) the *rate* of learning, whether it is a gradual evolution and combining of successive steps, or a sudden grasp of the problem as a whole. In general, mechanistic theories of learning have emphasized the random nature of the trials and the gradual development of learning through repeated attempts. Such theories are usually called *trial-and-error* explanations. In contrast, most psychologists who favor

* Accepted for publication December 2, 1937.

[1] Joseph Peterson, Studies in the comparative abilities of whites and negroes, *Ment. Meas. Monog.*, 1929, (no. 5), 1-156.

[2] C. E. Seashore and R. H. Seashore, *Elementary Experiments in Psychology*, 1935, 166-169.

[3] T. W. Cook, Amount of material and difficulty of problem solving: II. The disc transfer problem, *J. Exper. Psychol.*, 20, 1937, 288.

purposive interpretations of behavior have emphasized the appropriateness of the responses and maintain that the learning, when it does occur, is a sudden *insight,* usually involving a grasp of the problem as a whole.

The experiment outlined here illustrates the meaning of the two concepts, *trial-and-error* and *insight* by presenting a problem which theoretically can be solved by either method, and asking *O* to record his results and methods of learning. If the trial-and-error procedure is involved, the problem should require a considerable number of attempts before success is attained, and these attempts should become gradually more and more accurate. Such a procedure obviously requires too many trials to be very efficient, especially if any other methods are available. If on the contrary, the student is on the constant lookout for significant relationships, the learning should involve one or only a few trials for each problem, each trial being directed according to some plan in order to accomplish a certain object. If the foresight is concerned with the completion of some pattern of steps within the whole problem, this would be called *partial insight.* Such partial insights often consist of a series of moves which have previously been tried out by visual imagination. Full insight usually includes also a verbal formulation of a rule for making successive moves.

If, in addition to recording his successive steps in solving the problem, *O* treats the matter systematically, the experiment will also represent a good example of the method of inductive reasoning, in which he collects and records the data, analyzes and classifies it according to its major variables (*e.g.* blocks and positions) and examines each variable by itself and in combination for significant relationships which warrant a general conclusion. For difficult problems, this procedure should be the most efficient of all in saving time as well as guaranteeing the fullest understanding of the problem.

*Problem.* The puzzle of the pyramids. Mark three distinct positions, A, B, and C, across the top of a sheet of paper. Lay a series of disks on A in order of size so as to form a cone.[4] The problem is to transfer them in the same arrangement to B or C, *moving only one disk at a time* and *never laying a larger upon a smaller one;* sample problems are given below.

*Experiment 1: Three disks.*

*(A)* Transfer three disks from A to either B or C in 7 moves. In recording always

[4] For permanent equipment make disks or blocks of wood of various sizes. If these are made with a center hole to fit over pegs mounted on a wooden base, they will keep aligned and will not fall from their positions when placed. For group demonstration, different sized coins (penny, nickle, quarter, half-dollar, dollar) or squares of paper will suffice. If paper squares are used, a corner turned-up will greatly facilitate their handling.

label the smallest disk 1, the next smallest 2, etc. In moving disk 1 from A to B, simply record 1B, giving the number of the disk and the place to which it is moved. A correct solution for this problem is given below. Trace through it by actually making the moves with your disks, *e.g.* 1B, 2C, 1C, 3B, 1A, 2B, 1B.[5] This result has also been entered in the appropriate portion of Table I, which is a classification of *correct* results designed to aid in insightful learning. Notice that the table also gives solutions for the puzzle when only one disk or two disks are involved, and also for ending on either B or C as the designated goal. To avoid laborious trial-and-error procedure keep a constant lookout for significant relationships between the sequences of moves on the various problems. This problem involves so many related principles that almost anyone should be able to discover one or more of them during the experiment. As soon as one is discovered, write it down immediately after the series of moves, and, similarly, if it is later discovered to be inadequate, note this fact at the point where it was discovered. Record also by underlining whether the moves seem to be random (R), partial insight (PI), or a full insight (I), according to some rule.

TABLE I

CORRECT SOLUTIONS OF PYRAMID PUZZLE

Watch for repeated patterns! Upon completing this table, review the list of principles discovered. If time permits verify each one and arrange them in some logical order in the space below. These principles are summarized in more complete form at the end of the article as an illustration of an inductive solution of the problem.

| No. of disks | Series of moves from A to B | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 1B | | | | | | | | | | | | | | | |
| 2 | 1C | 2B | 1C | | | | | | | | | | | | | |
| 3 | 1B | 2C | 1C | 3B | 1A | 2B | 1B | | | | | | | | | |
| 4 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | |
| 5 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — |
| | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | |

| No. of disks | Series of Moves from A to C | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 1C | | | | | | | | | | | | | | | |
| 2 | 1B | 2C | 1C | | | | | | | | | | | | | |
| 3 | — | — | — | — | — | — | — | | | | | | | | | |
| 4 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | |
| 5 | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — |
| | — | — | — | — | — | — | — | — | — | — | — | — | — | — | — | |

*(B)* Transfer three disks from A to C in 7 moves. Record (1) each move as it is made, (2) the nature of the trials: (R), (PI) or (I), (3) the principles discovered, and (4) enter correct series of moves in Table I. Start a new line every time a mistake is discovered. Use the margins of the page if necessary. Follow the same procedure for all further experiments.

*Experiment II. Four disks.*

*(A)* Transfer four disks from A to either B or C in 15 moves, recording as before. (For actual use, construct a mimeographed sheet having all directions, with spaces for recording each trial as in Experiment I, listing each successive step, *e.g.* Experi-

[5] In recording, always start at the left margin and keep records in a neat tabular form.

ment II A, II B, etc. Provision of materials for recording is of course essential for any worthwhile experiment.

*(B)* Transfer four disks from A to the opposite position from that done in the previous experiment, in 15 moves.

*Experiment III. Five disks.*

*(A)* Transfer five disks from A to either B or C in 31 moves. (Arrange recording blank to have 16 blanks in top row, 15 in bottom row, as this brings out an essential relationship.

*(B)* Transfer five disks from A to the opposite position from that in the previous experiment in 31 moves.

The problem may be continued for as many disks as desired, and with any setting so as to emphasize various factors.

A review of a way in which the problem meets the major requirements of experimental material will suggest some of its possible uses in psychology.

*(1)* Being simple in construction and operation and having only two simple rules, there are few artificial restrictions upon the method of learning which may be employed. It is thus a fair test of the relevancy of trial-and-error *vs.* insight theories of human learning. By simply adding a few directions, such as those for tabulating all moves, insight may be facilitated, and by a few further directions inductive analysis may be called into play.

*(2)* It is suitable for the most intensive individual analysis or for extensive group measurements. It avoids the disadvantages of most mechanical puzzles and the T-F-L puzzles which depend upon the discovery of one major principle by presenting at least 50 coördinated principles, any one or a few of which are also useful singly.

*(3)* The range of difficulty may be varied from that of pre-school children to that of the most highly trained adults by simply changing the number of disks or by limiting the goal to one instead of two alternative end positions. Everyone can solve a number of levels of the problem within a short time, and quantitative differences may be stated in terms of time, moves, or number of principles discovered.

*(4)* Each move may be recorded by simply giving the number of the disk and the position to which it is moved, *e.g.* 1B.

*(5)* The process of problem solving may be varied or repeated by simply increasing the difficulty on successive trials, thus verifying earlier findings. Fractionation of the problem is likewise facilitated by its stepwise arrangement.

*(6)* Its motivational value is high. It is interesting, and has at times been used as a party entertainment.

*(7)* Accuracy is checked at any stage by simply comparing the number of moves with the known minimum for each level of difficulty.

*(8)* No difficult motor manipulation or technical knowledge is necessary.

Preliminary work on this problem yields the following results. (1) Solutions practically never begin with moves which are entirely random. They are at least limited to those having some degree of relevance. (2) Trials (or moves) may be either implicit (imagined) or explicit but both

must be recorded in any complete description. (3) An insight is only one of many trials, and even an error may be an imperfect insight. (4) Insights are of at least three different kinds: (a) a vague hunch that a certain move or series is correct, although no reason can be cited, (b) clear imagery such as visualization of a series of moves and their consequences, and (c) discovery of verbal principles or rules. All three types may be either partial or complete. The term insight thus acquires a much broader and more intelligible meaning. (5) At the college level the supposed relationship between high intelligence test scores and the tendency to employ rational as opposed to trial and error methods is not as close as one might expect.

Methods of approach do, however, vary significantly with changes in the difficulty of the problem. Rapid trial-and-error is sufficient for simple stages. Searching for informal insights suffices for stages of moderate difficulty, but after eight years of class use of these methods only four or five principles had been discovered, whereas a few months of inductive analysis of more difficult stages of the problem brought to light about fifty principles. This analysis showed that there are three major variables, (a) the order in which the disks move, (b) the order of the positions to which they move, and (c) the goal specified. The fifty principles all fall under these three classifications and a fourth which combines them into one master equation which predicts all moves for any number of disks or specified goal. The application of these principles for verification and prediction of specific problems is a further possibility for specific training.

*Inductive reasoning.* The experiment up to this point has been designed to illustrate and clarify the concepts of insight and trial-and-error as they apply to methods of learning in simple problems. For longer and more difficult problems as in scientific research, where more accurate and economical methods are necessary, a third and more systematic method is available; namely, *inductive reasoning.*[6] This makes use of both the previous principles but in a more orderly and planned way.

The main steps of inductive reasoning may be easily demonstrated or tested by a further treatment of the results already obtained in the same pyramid puzzle problem, thus taking advantage of the interest and knowledge already gained. Having collected, recorded, and classified the data in Table I, the results may next be *examined for any constant tendencies or principles.*

[6] This part of the experiment may be treated separately or at a later time.

In any large and difficult problem, one of the first steps is to analyze out its main variables, which may then be attacked one at a time, in simpler fashion. A careful inspection of the pyramid puzzle shows that the main variables are: (1) the order of size in which the various blocks move, (2) the order of positions to which each moves, and (3) the effect of changing goal positions. Table I was constructed to isolate each of these factors, to bring out its *variations* from simplest to most complex examples, as well as the advantage of *repetition* of observations because of the fact that simpler problems are repeated as the first stages of each succeeding more complex problem.

A further reason for constructing Table I, a list of correct moves only, was to *classify* the observations, thereby emphasizing one variable at a time, and bringing out more clearly the significant aspects of principles to be discovered by discarding the misleading records of incorrect moves. The accumulation of large numbers of notes or records tends to increase the difficulty of clear interpretations, making it necessary to *summarize* the principal trends in tables, graphs, or indices. Notice in Table I the many remarkable examples of symmetry in recurring patterns and slight deviations from exact regularity. The very name, *pyramid* puzzle, should suggest to those who have any interest or training in mathematics that its series should be highly symmetrical and regular. The table itself is symmetrical, the order of blocks moved is symmetrical, the order of positions to which each block moves is regular, and the change in positions according to the particular goal designated is regular. If time is available, many regularities of this sort may be detected and stated as accurately as possible, checking the principle by applying it to similar or slightly more difficult problems. The first recognitions of these principles are apt to be partial insights, which are later refined by trial-and-error to the status of complete insights for each part of the problem. A complete discovery of all the principles involved would take much longer, but after trying a few the investigator may be interested to compare these with the rather complete list worked out by the writer as an example of the types of things which may be discovered in such situations.

The principles stated in the following summary have been simplified and arranged in order from a much larger list of discoveries made by students and instructors who have tried the problem. While it is probable that they may be still further clarified we may feel sure that we have here the most essential ones because the classification agrees with the list of variables and furthermore (the test of synthesis) the essential principles dealing with each variable may be combined in the one master equation for

predicting every phase of the complete list of moves for any problem.

A few further points about the inductive method. The average number of principles discovered by students who did not keep a record similar to Table I was only about 1.5 per student. The use of a planned record should be more productive because of the more systematic classification of results, which favors noticing recurrent patterns.

PRINCIPLES INVOLVED IN THE SOLUTION OF THE PYRAMID PUZZLE

*I.* Principles describing numbers and orders of disk moves (toward the designated goal, *e.g.* A to B.)

*(A)* In examining Table I it is seen that all vertical columns involve the same disk (but not the same position).

*(B)* All trials, regardless of number of disks, which move toward the same final goal, start out by moving each disk in the same order (but not to the same position) as in the trial for the next smaller number of disks. The middle move is always the one involving the largest disk and the remaining moves are in the same order (but not to the same positions) as for the first half of the series. The equation for this is given P, N, P where P = series of disks moved for next smaller (previous) number of disks, and N = new or largest block.

*(C)* All series require an *odd* number of moves, as may be inferred from the above explanation.

*(D)* The number of moves in any problem is equal to twice the number for the next smaller series of disks plus 1.

*(E)* For any given number of disks, the bottom disk moves once, the next smaller disk twice and each successively smaller disk twice as many times as the next larger one.

*(F)* The smallest disk always moves as many times as all the other disks combined.

*(G)* The smallest disk starts with the first move and continues to move every alternate time. Disk 2 starts with the second move and continues to move every fourth time, etc. The formula for the moves is as follows: Disk 1 starts with the first move and continues to move every $2^n$ times (where n equals the number of the block). All succeeding disks start with the $2^{n-1}$ move in the series and continue to move every $2^n$ time.

*II.* Principles describing positions of moves in the series (toward the same designated goal, *e.g.* A to B).

*(A)* in Table I each vertical column contains an alternation of B and C moves, but A remains constant. (Except for the last line, which is out of position due to lack of space.)

*(B)* The *positions* for each move in a new series may be derived by the following principles:

*(1)* For the first half of the series, reverse positions B and C (keeping A constant) from the series for one less disk, and then move the largest disk to the designated goal.

*(2)* For the last half of the series, reverse positions A and C (keeping B constant) except that the last two moves will always be disks 2 and 1 to the goal.

*(C)* The last two moves of all problems that end on the same goal will always be the same in respect to the numbers of the disks and the position to which they are moved.

*III.* Principles involving difference in designated goals.

*(A)* When two problems involve the same number of disks but different designated goals, the order of moves will be the same, but the positions of B and C will be reversed, A moves remaining constant.

*IV.* Principles dealing with interrelations of variables I, II, and III.

*(A)* When the designated goal is the same, problems, involving odd numbers of disks begin by moving the first disk to the designated goal, while even numbers of disks begin by moving the first disk to the undesignated position. This principle applies also to patterns within the complete series of moves.[7]

*(B)* If we call any movement from A to C or vice versa a "skip" and any movement to an adjacent pile a "move," then for any problem with a designated goal C and an odd number of disks there will be a cycle "skip one disk, and move two disks;" this cycle being repeated all through the series. When the designated goal is C and the number of disks is even, the rule is "move two, skip one." When the designated goal is B the rules are just reversed from those just stated (for odd and even-numbers of disks).

*(C)* Cycles of positions for each Block.

| | | | | |
|---|---|---|---|---|
| Block 1— | C | B | A | Reverse alphabetical cycle |
| Block 2— | A | B | C | Forward alphabetical cycle |
| Block 3— | C | B | A | Reverse alphabetical cycle |
| Block 4— | B | C | A | Forward alphabetical cycle |

Odd-numbered disks reverse, even-numbered disks forward, for Goal C; just the opposite for Goal B.

*(D)* Cycles of positions for pyramids. Always seek to uncover the next larger disk from the original position, making a vacant space for it by piling all smaller disks in a pyramid. In the first half of the problem build all even-numbered pyramids on the undesignated (U) position, and all odd-numbered pyramids on the designated (D) position.

*(E)* With two exceptions, both of which are predictable, all series of positions tend to fall into symmetrical patterns of fours, *e.g.* BCCB, in which each new letter follows the forward order ABC both within the pattern and the reverse from one pattern to the next.[8] These patterns of four moves are further grouped into patterns of three groups each, and then repeat in the same order.

*(1)* As an illustration of the exceptions, the series of moves for a six-disk problem is: BCCB ABBC CAAC BCCB ABBA CAAB BCCB ABBA CAAC BCCA ABBA CAAC BCCB ABBC CAAC BCC, in which the exceptions have been underlined. Further inspection reveals

*(a)* that exceptions occur when an even-numbered disk of size four or higher is moved *(i)* by changing to the next letter in the forward series ABC, and *(ii)* only at final positions, on numbers of moves divisible by 8;

*(b)* that the final pattern always omits the fourth position; and

---

[7] This is the commonest and most useful discovery. Often, however, it is discovered only in partial form.

[8] See point E-2 below.

*(c)* that the last two moves are to the designated goal.

*(2)* For even-numbered problems A-C and odd-numbered problems A-B, the series will be CBBC ACCB BAAB. For odd-numbered problems A-C and even-numbered problems A-B, the series will be BCCB ABBC CAAC.

*(F)* The formulae for writing the order of moves by disk and position for odd-numbered problems are as follows:

| disk ..... | first | any even-numbered | any odd-numbered |
|---|---|---|---|
| number of 1st move ..... | 1 | $2^{n-1}$ | $2^{n-1}$ |
| first position of disk ..... | D | U | D |
| cycle ..... | R | F | R |
| no. moves to recurrence ..... | $2^n$ | $2^n$ | $2^n$ |

where n = the number of the disk
D = designated position
U = undesignated position
F = forward cycle UDS
R = reverse cycle DUS
S = starting position

For even-numbered problems, starting positions D and U given above are reversed.

The number of moves required for solution for N number of disks is $2^N$—1.

## APPARATUS NOTES

### An Inexpensive Stop-Clock

The electrical alarm clock made by the Westclox Company for the household may readily be used as a stop-clock, and when so used it is more reliable than the ordinary stop-watches that are usually employed in the administering of psychological tests.

The knob at the back of the clock is used to start and to stop the timing-mechanism. Before using the clock for timing, it is advisable to stop the second hand at a point about 5 units before the zero position. When ready to begin timing, start the clock by spinning the starting-knob, and say "go" just as the second hand passes the zero position. This procedure corrects for 'racing' and insures an even start. At the end of the interval to be timed, the knob is rotated quickly to stop the clock, and released gently to prevent the clock from starting. Time may then be read in hours, minutes, seconds, and estimated fractions of seconds when the pointer stops between points of calibration. This model seemed to be more nearly free from the irregular speeds of the seconds pointer, a defect found in several other models.

Lehigh University ADELBERT FORD

# NOTES AND DISCUSSIONS

## Recent Advances in Statistical Theory and Applications

The field of statistics is developing so rapidly and new applications of old techniques are being made at such a rate that it is impossible to do justice to all workers. Many of the technical journals were not examined, and it is too much to hope that, in the journals examined, no important study has been overlooked. The material presented has been selected for one or more of the following reasons: importance, general interest, availability, and application to problems of psychology.

*Correlation.* Our present statistical techniques go back, for the most part, to Gauss' paper, *Theoria Combinationis Observatum,* in which he proposes that the square of the deviation from the mean be used rather than the absolute value. This is the familiar $\Sigma x^2$ term in the standard deviation, analagous to the $\Sigma xy$ term in the correlation coefficient. Laplace, with whose work Gauss was familiar, chose to work with the absolute values of the deviations. The use of these values eliminates the possibility of a simple analytical treatment, and for this reason modern mathematical statisticians have ignored Laplace's method and chosen the method of squared deviations. Lindeberg,[1] however, has been able to derive an expression of the standard deviation of a measure of skewness on Laplacian lines. His success caused Bernstein to approach the problem of correlation and regression in terms of linear expressions. This paper,[2] is one of the most interesting and probably most important papers to appear in 1937. The choice of title seems unfortunate, for it in no way indicates the importance of the contents. Actually, the paper presents a method of securing correlation and regression coefficients by means of linear expressions, in other words, in terms of absolute deviations. These new functions are much easier to compute than the Pearson r, and three methods of computation are given by Bernstein.

In the case of a normal bivariate distribution, Bernstein's new expressions give less information than do Pearson's functions. For other types of distributions, however, this is not generally true. These new formulas are particularly valuable in those cases where the extreme measurements are unreliable, because undue weight is not given by squaring or multiplying the deviations. There is no need to point out applications, for all experimental psychologists have at times been confronted with the problem of what to do with extreme but unreliable measurements. Since the new expressions are linear, no troublesome adjustments for grouping are necessary. The simplicity of the formulas is apparent from an examination of the regression coefficients. It should be noted that only partial sums are needed for these coefficients. These formulas deserve serious consideration by psychologists whose data are not strictly normal, or whose data contain extreme deviations as a result of errors of measurement.

---

* Read before the Eastern Psychological Association, Saturday, April 2, 1938.

[1] J. W. Lindeberg, Über die Begriffe Schiefheit und Excess in der mathematischen Statistik, *Skand. Aktuarietidskr.,* 8, 1925, 106-127.

[2] F. Bernstein, Regression and correlation evaluated by a method of partial sums, *Ann. Math. Stat.,* 8, 1937, 77-89.

The old problem, correlation of sums, has been the subject of some recent work. Conrad has developed an appropriate formula for determining the correlation between a single element of a composite and the remainder of the composite.[3] This function will be found useful in an item analysis to determine the correlation between an item and the rest of the test. It can also be applied to the problem of the correlation between a subtest and the remainder of the battery. A closely related problem is the determination of the average correlation between one variable and the N—1 other variables being studied. Edgerton has developed such a formula which has the advantage that the desired value can be found without computing the various intercorrelations.[4]

A formula for the intercorrelation among multi-scores, such as are secured from interest and personality blanks, has been developed by Royer.[5] This formula is used for estimating the amount of correlation due to the relationships among the keys, and is well worth the attention of those working with questionnaires.

The possibility of accounting for factor loadings in genetic terms has been contemplated by many psychologists. In this connection it is of interest to note that Price[6] using formulas developed by Wright for the purpose of determining correlation between mates for a given somatic character,[7] has pointed out the effect of homogamy on the intercorrelation of capacity traits. Bartlett,[8] following Price's lead, has examined correlations obtained in equilibrium and studied their relations to various theories of mental abilities.

The Pearson correlation coefficient is often used to estimate the relationship between scores on a test in standard time with scores under unlimited time conditions. Odoroff[9] has pointed out that the Pearson $r$ gives too high a value and suggests that in such studies a formula derived by Harris[10] for the elimination of spurious effects might well be applied.

One of the most troublesome and at the same time common statistical problems confronting psychologists is what to do about curvilinear regression. It is well known that most psychological measurements are made with an arbitrary scale, but what is not so generally recognized is the fact that some other scale might just as well be used. There is nothing sacred about the units of any scale. The raw units may be used, or any function of them that is convenient. Moore clearly recognized this

[3] H. S. Conrad, On the calculation of the correlation between a single element of a composite and the remainder of the composite, *J. Educ. Psychol.*, 26, 1935, 611-615.

[4] H. A. Edgerton, A formula for finding the average correlation of any one variable with the (n—1) other variables without solving any of the individual correlations, *ibid.*, 26, 1935, 373-376.

[5] E. B. Royer, A formula for intercorrelations among multi-scores, *ibid.*, 27, 1936, 457-466.

[6] B. Price, Homogamy and the intercorrelation of capacity traits, *Ann. Eugen.*, 7, 1936, 22-27.

[7] S. Wright, Systems of mating: III. Assortative mating based on somatic resemblance, *Genetics*, 6, 1921, 144-161.

[8] M. S. Bartlett, Note on the development of correlations among genetic components of ability, *Ann. Eugen.*, 7, 1937, 299-302.

[9] M. E. Odoroff, A correlational method applicable to the study of the time factor in intelligence tests, *J. Educ. Psychol.*, 26, 1935, 307-311.

[10] J. A. Harris, The correlation of a variable and the deviation of a dependent variable from its probable value, *Biometrika*, 6, 1909, 438-443.

and illustrates a number of methods for securing linear relationship where the original units result in non-linear relations.[11] For those confronted with this type of problem, time spent with Moore's article will be well repaid.

Another problem that is constantly arising is how to combine correlations obtained from several groups into a single coefficient. It is well known that averaging correlations is a hazardous procedure, but it is not so well known that it is possible to combine two positive correlations of, say, 0.80 and get a high negative correlation. It may be desirable to synthesize correlations based on one's own data or from data found in literature. If the correlations, means, standard deviations, and N's are known, it is a simple matter to evaluate correctly the desired coefficient by a formula given by Dunlap.[12]

The problem of correlation cannot be left without considering recent work on prediction. In many investigations, the psychologist is confronted with the problem of selecting one or more criteria from a number of possible criteria. For example, in the prediction of college success, one criterion might be grades, another popularity, and a third, salary in the initial job. One possibility is to select as the criterion not one of these, but the linear function of the criteria whose means square correlation with the independent variable is a maximum. Hotelling has developed a method for combining various criteria into a single composite,[13] which should be most valuable to those workers confronted with the problem of choosing one criterion from a multiplicity of criteria. Kurtz has approached this problem somewhat differently.[14] He argues that we may wish to use the various criteria successively, as, for example, in guidance various criteria may need to be considered separately. He determined a simple method for securing a unique set of weights for the independent variables, such that the average of the correlations with several criteria will be a maximum. The choice between this and Hotelling's method will depend on what the worker desires.

Rank correlation generally is thought of as an easily calculable substitute for the product moment correlation. It is true, it is easy to calculate, but it has other properties that are far more valuable. Hotelling has for years emphasized this statistical Cinderella, but only recently has rank correlation commenced to receive general recognition. Hotelling and Pabst point out that it can be used as a test of the existence of correlation without any assumption as to the type of distribution, either normal or skew, and when used as such a test, is capable of exact interpretation in terms of probability.[15] Furthermore, the rank correlation often can be combined with the method of contingency to secure a rigorous test of relationship. Such a combination is permissible whenever it can be shown that the two measures are independent. An excellent example is given in a hypothetical study of the sex differences in grades assigned in a particular subject. Testing the distribution by the chi

[11] T. V. Moore, The reduction of data showing non-linear regression for correlation by the ordinary product-moment formula; and the measurement of error due to curvilinear regression, *J. Educ. Psychol.*, 28, 1937, 205-214.

[12] J. W. Dunlap, Combinative properties of correlation coefficients, *J. Exper. Educ.*, 5, 1937, 286-288.

[13] H. Hotelling, Relations between two sets of variates, *Biometrika*, 28, 1936, 221-277.

[14] A. K. Kurtz The simultaneous prediction of any number of criteria by use of a unique set of weights, *Psychometrika*, 2, 1937, 95-101.

[15] H. Hotelling, and M. R. Pabst, Rank correlation and tests of significance involving no assumption of normality, *Ann. Math. Stat.*, 7, 1936, 29-43.

square method or by rank correlation did not reveal a significant difference between the sexes, but when the two measures were combined, the difference was shown to be significant. Obviously, if either test reveals a significant difference, it is not necessary to apply the other. Analagous problems are constantly occuring, so that this method probably will develop into a standard procedure for testing significance.

*Analysis of variance.* The use of ranks is not limited to the problem of correlation. In many psychological and social problems, it is impossible to assume a normal distribution of the variable under consideration, and in many cases only ordinal information is available. In such problems, the application of the analysis of variance technique to isolate the variability due to different factors is not justifiable. Friedman,[16] recognizing that in ranked data it is not necessary to assume normality, has developed a chi square technique for such problems. The method uses only information on order. From the computational standpoint, it is extremely simple, requiring about one fourth the time needed to analyze the variance. For normal distributions, the method is not as efficient as analysis of variance, but should be used when normality cannot be assumed, or the problem of computational time is important.

Fertig has shown that the analysis of paired samples can be successfully treated by the method of analysis of variance.[17] He has demonstrated that the interaction variance is algebraically equivalent to the standard error of the difference between means. When there is no correlation between the variates, then the interaction variance is a valid and unbiased estimate of the experimental error. The use of the interaction variance circumvents the objection that in the difference method some of the variation not properly ascribable to experimental error is included. When the two series are positively correlated, both methods should be used and, if either shows a significant difference, cognizance should be taken of it.

The method of analysis of variance is especially important in those problems where the investigator wishes to determine the significance of trends in seasonal variation and to eliminate or control such trends. The method is strongly recommended to those whose problems involve a time variation such as occurs in learning problems, or in successive measurements to show shifts of attitudes, or in repeated measurements to show progressive mental deterioration, or in any problem where one function is varied systematically, as, say, where increasing quantities of a drug are administered successively, and measurements taken of some variable. An excellent example of the application to seasonal variation problems is that of Jellinek and Looney who studied the seasonal variation of blood cholesterol.[18] The writer, in an unpublished study of the seasonal variation of blood sugar content of rabbits applied this technique with satisfactory results.

*Reliability.* Jellinek has approached the problem of reliability with a fresh point of view.[19] He points out that a self-correlation of $+1$ does not mean individuals

[16] M. Friedman, The use of ranks to avoid the assumption of normality implicit in the analysis of variance, *J. Amer. Stat. Assoc.*, 32, 1937, 675-701.

[17] J. W. Fertig, The use of interaction in the removal of correlated variation, *Biom. Bull.*, 1, 1936, 1-14.

[18] E. M. Jellinek, and J. M. Looney, Studies in seasonal variation of physiological functions: I. The seasonal variation of blood cholesterol, *ibid.*, 1, 1936, 83-95.

[19] E. M. Jellinek, Measurements of consistency of fasting oxygen consumption rates in schizophrenic patients and normal patients, *ibid.*, 1, 1936, 15-45.

produced identical results on repeated tests. This failure of the product-moment coefficient to take account of any systematic shift in an individual's scores due to change of technique, learning, familiarity, or practice, he considers a serious fault. If what is wanted is a measure of level maintained, that is, consistency of results, then the intra-class correlation should be used. Excellent as this method is, Jellinek feels that an even better way of expressing consistency is in terms of the average intra-set variance. Estimates of the inter-individual and intra-individual variances are obtained, and a measure of consistency found by dividing the inter-individual variance by the intra-individual variance, which can be tested by Snedecor's "F" function. An advantage of this method is that if the inter-individual variance is significantly greater than the intra-individual variance, there is evidence that the population studied is heterogeneous. This test may also be applied to two intra-individual variances to see if the individuals in one sample show greater consistency in a given variable than individuals of another sample. For example, if it is desirable to determine whether delinquents are more or less consistent in a given experience, trait, or function than a control group, this ratio provides a very sensitive test.

Thouless has pointed out that scores secured from two applications of a test may differ due to unreliability of the test, or because the function measured has undergone real change.[20] He proposes to test for function variation by a double test-retest technique, where the difference between $A_1$ and $A_2$ is correlated with the difference between $B_1$ and $B_2$, where the letters indicate the forms, and the subscripts the dates of the tests. If function fluctuation is not present, then this correlation will be zero within the sampling errors. Those dealing with measures of attitudes, interests, temperament and other functions will do well to consider Thouless' index when reporting reliability.

It has been urged by many workers that tests and measurements should be repeated at stated intervals, and that emphasis should be placed on progress, gains in achievement, increases in ability, and other dynamic concepts. If this concept is to be employed, it is necessary to have a measure of the reliability of the gains. Shuttleworth has derived a formula for this purpose.[21] Examination of this function reveals that, in general, the reliability of the increments is lower than the reliability of the original measurements, becoming zero when the correlation between the initial and final measurements is as high as the reliability of the original scores.

An interesting point is the contention of Stephenson that in choosing between two tests for a battery, the test having the greater saturation with "g" should be selected, even though it has the lower reliability.[22] The application of this concept to test construction is so obvious as to need no further comment.

*Sampling theory.* Many problems in psychology, physiology, and psychiatry deal with variables that are highly interrelated and cannot easily or adequately be treated separately. Often it is impractical, or even impossible to analyze the separate effects,

---

[20] R. H. Thouless, Test unreliability and function fluctuation, *Brit. J. Psychol.*, 26, 1936, 325-343.

[21] F. K. Shuttleworth, The reliability of increments, *J. Educ. Psychol.*, 26, 1935, 312-313.

[22] W. Stephenson, Factorizing the reliability coefficient, *Brit. J. Psychol.*, 25, 1934, 211-216.

and in such cases the simultaneous treatment of the whole is necessary. Fertig,[23] extending the work of Neyman and Pearson[24] and that of Wilks[25] and Pearson[26] on the lambda function for testing hypotheses regarding the origin of one or more observed samples, has developed a method for estimating whether the sample as a whole meets a specified hypothesis, thus eliminating the question as to whether the hypothesis holds for the separate statistics.

The rationale of testing a hypothesis consists in arranging an experiment so that the probability of obtaining a particular set of results can be calculated. If the actual results are such that the probability is excessively small, then the hypothesis is considered false. Unfortunately, in many experiments, the probability is not small, so it is of utmost importance that criteria be developed to test such negative results. Dorfman has developed three such criteria, one for determining when it is unprofitable to make more observations, one for estimating the reliability of presumptive conclusions, and one for determining the maximum amount of error which may result if the null hypothesis is taken as a working basis.[27]

Madow has developed a set of theorems for the statistical examination of hypotheses which are false if any one of several things are true.[28] Essentially, he reduces the distributions by which the hypotheses are to be tested to a joint distribution of the statistics which contain the information offered by the data as to the truth or falsity of the things which will negate the hypothesis.

The widespread use of the method of matched groups, which essentially is an attempt to apply "the law of the single variable" makes any new statistical development on this controversial problem a matter of major importance. Johnson and Neyman generalize this problem in such a way that a more detailed analysis of the data is secured, and with greater accuracy of results.[29] They also introduce in this paper the idea of "regions of significance" which can be applied to many problems in educational and psychological research.

Baker has studied a related problem;[30] namely, what is the probability that the mean of a second sample will lie within a certain multiple of the standard deviation of the first sample from the mean of the first sample? His results have the definite advantage of allowing the worker to express differences in terms of the unit in which they were first expressed and to specify definitely the point from which the measure-

[23] J. W. Fertig, The testing of certain hypotheses by means of lambda criteria with particular reference to physiological research: I. The drawing of one or more samples from completely or partially specified populations, *Biom. Bull.*, 1, 1936, 45-82.

[24] J. Neyman and E. S. Pearson, On the use and interpretation of certain test criteria for purposes of statistical inference, *Biometrika*, 20, 1928, 262-294.

[25] S. S. Wilks, Certain generalizations in the analysis of variance, *ibid.*, 24, 1932, 471-494.

[26] E. S. Pearson and S. S. Wilks, Methods of statistical analysis appropriate for k samples of two variables, *ibid.*, 25, 1933, 351-378.

[27] R. Dorfman, The use of fiducial probability in the interpretation of inconclusive experiments, *Biom. Bull.*, 1, 1937, 109-128.

[28] W. G. Madow, Contributions to the theory of comparative statistical analysis: I. Fundamental theorems of comparative analysis, *Ann. Math. Stat.*, 8, 1937, 159-176.

[29] P. O. Johnson and J. Neyman, Tests of certain linear hypotheses and their application to some educational problems, *Stat. Res. Mem.*, 1, 1936, 57-93.

[30] G. A. Baker, The probability that the mean of a second sample will differ from the mean of a first sample by less than a certain multiple of the standard deviation of the first sample, *Ann. Math. Stat.*, 6, 1935, 197-201.

ment is made. Many differences which formerly would have been considered significant may not be considered so by Baker's method.

In connection with sampling theory, two valuable tables have been published recently. Hendricks[31] has developed further tables for Student's function. Edgerton and Starbuck[32] have published a table for use with the computation of Fisher's "t" function. This group of papers materially enhances the range of problems susceptible to statistical analysis.

*Nomographs.* The possibility of reducing the labor of many calculations by means of nomographs is being recognized more and more. Most of the recent nomographs are adapted to various problems in item analysis or in test construction. Arnold has developed a nomograph for estimating the validity of test items,[33] and Zubin has a graphic method designed for the same purpose.[34] A nomograph for estimating the validity coefficient of a lengthened test has been prepared by Voss,[35] and Dunlap[36] has developed one for evaluating biserial r, while Kuder[37] presents in a single article nomographs for point biserial r, biserial r, and fourfold correlations. Three nomographs for use with the Spearman-Brown formula are given by Arnold and Dunlap.[38] They are (a) for estimating the correlation when the test is lengthened 'N' times, (b) for the standard error of a correlation estimated by the Spearman-Brown formula, and (c) for the standard error of 'N.' The only nomograph developed for the use of the clinical psychologist is one for determining the Heinis Personal Constant, presented by Voss and Dunlap.[39]

*Statistics of matchings.* In certain studies of personality, subjects have been required to match samples of handwriting against character sketches of the same individuals. The number of correct matchings is then compared with the number expected by chance. Chapman[40] has discussed the assumptions underlying such comparisons and points out that Salvosa's tables[41] furnish the desired probabilities. In a later study[42]

[31] W. A. Hendricks, An approximation to "Student's" distribution, *ibid.*, 7, 1936, 210-221.

[32] H. A. Edgerton and E. O. Starbuck, A table to aid in the computation of Fisher's "t" function for comparison of two means, *Psychometrika*, 2, 1937, 77-84.

[33] J. N. Arnold, Nomogram for determining validity of test items, *J. Educ. Psychol.*, 26, 1935, 151-153.

[34] J. Zubin, Note on a graphic method for determining the significance of the difference between group frequencies, *ibid.*, 27, 1936, 431-444.

[35] H. A. Voss, A nomograph for estimating the validity coefficient of a lengthened test, *ibid.*, 5, 1937, 290-291.

[36] J. W. Dunlap, Nomograph for computing bi-serial correlations, *Psychometrika*, 1, 1936, 59-60.

[37] G. F. Kuder, Nomograph for point biserial r, biserial r, and fourfold correlations, *ibid.*, 2, 1937, 135-138.

[38] J. N. Arnold and J. W. Dunlap, Nomographs concerning the Spearman-Brown formula and related functions, *J. Educ. Psychol.*, 27, 1936, 371-374.

[39] H. A. Voss and J. W. Dunlap, Determination of the Heinis personal constant, *J. Consult. Psychol.*, 1, 1937, 57-61.

[40] D. W. Chapman, The statistics of the method of correct matchings, this JOURNAL, 46, 1934, 287-298.

[41] L. R. Salvosa, Tables of Pearson's Type III function, *Ann. Math. Stat.*, 1, 1930, 191-198.

[42] D. W. Chapman, The generalized problem of correct matchings, *ibid.*, 6, 1935, 85-95.

he examines the generalized problem of correct matchings and develops estimates of significance for the occurrence of a given number of correct matchings and for the mean number of correct matchings produced by a group of individuals working independently but with the same material. Chapman has also developed a formula for determining an individual's score when he matches N items against M others.[43]

Zubin[44] has advocated a mean square contingency coefficient for evaluating matched data, while Vernon[45] points out that the contingency coefficient depends upon the deviations from chance frequency in all cells and not only upon the deviations of the correct responses. Vernon proposes to modify the method slightly and obtains highly satisfactory results when the modified method is checked experimentally, and against Chapman's tables.

In a later study Vernon[46] points out that the contingency coefficient possesses two standard errors when used this way, one referring to the variation in the judges, and the other to the variation in the sets of material to be matched, and provides a formula for these standard errors. Rosander approached the problem from a different viewpoint and has developed a formula for determining the standard error of a mean rank order.[47]

*Test construction.* There has been for many years an unfortunate tendency on the part of the experimental psychologist to look down upon the "mental tester" and for the "tester" to sneer at the "brass instrument" psychologist. The problems of the two are not irreconcilable; in fact, it has long been suspected that many of them are identical. Guilford states that any psychophysical judgment may be regarded as a response to a test item if we are interested only in whether the subject renders a correct response.[48] Using the Seashore Tests, he has shown that the easiness of an item as determined by absolute scale methods is proportional to the logarithm of the magnitude of the stimulus. He proposes that the unit of absolute scaling as applied to test items may become a satisfactory unit for all "S" scales in the traditional psychophysical problems. This attempt to coördinate the technique of the two fields and develop a sympathetic understanding of the problems deserves serious consideration by both groups.

In measuring psychological values which can not be objectively evaluated, such as social attitudes, nationality preferences, and esthetic qualities, it is necessary to construct a scale, using one of the psychophysical methods, such as the Rank Order Method, the Method of Paired Comparisons, or the Method of Successive Intervals. The choice of scaling method is important unless the methods are linearly related. Saffir's experimental study of this problem clearly demonstrates that the methods are

---

[43] D. W. Chapman, The scoring of matching-tests with unequal series of items, *J. Educ. Psychol.*, 27, 1936, 368-370.

[44] J. Zubin, The chance element in matching tests, *ibid.*, 24, 1933, 674-681.

[45] P. E. Vernon, The evaluation of the matching method, *ibid.*, 27, 1936, 1-17.

[46] P. E. Vernon, A note on the standard error in the contingency matching technique, *ibid.*, 27, 1936, 704-709.

[47] A. C. Rosander, The standard error of a mean rank order, *ibid.*, 27, 1936, 193-196.

[48] J. P. Guilford, The psychophysics of mental test difficulty, *Psychometrika*, 2, 1937, 121-133.

linearly related, so the choice of method is one of practical convenience rather than one of relative validity.[49]

Horst emphasizes that the method of factor analysis is superior to the Method of Equal Appearing Intervals, and the Method of Constant Stimuli for scaling attitude items.[50] Since such scales do not measure a pure function, he proposes to use the first factor loading as the major function, and then, if desired, use the other loadings as measures of other functions. This use of factor analysis has been pointed out earlier by Thurstone.

The selection of items for a test has become an extremely technical problem during the past decade, but it has remained for Merrill to point out explicitly that underlying all item analysis is the assumption that the items actually do differ from each other with respect to difficulty and particularly with respect to validity.[51] If items are homogeneous with respect to validity, but, for a particular sample of individuals, vary in validity due to sampling error, then there is no need to select particular items. Merrill points out that the familiar chi square function supplies the test constructor with a means of testing a given set of items to see whether the validity coefficients vary more than would be expected by chance in a homogeneous distribuion.

A new index for item selection has been proposed by Sletto that, from a logical examination, appears valuable in attempts to construct "pure" tests.[52] The method merits an extended experimental comparison with the better current methods.

Two excellent studies on the difficulty of test items are those of Richardson[53] and Voss.[54] Richardson has shown that the difficulty of a test element is directly related to the maximum differential validity, while Voss studied the relationship between item difficulty and reliability and found a definite relation. Tests composed of items having approximately fifty percent difficulty give the highest reliability and validity. Items between the upper and lower quartiles in difficulty appear to be satisfactory for most purposes.

An interesting paper by Walker shows that the distribution of scores is a function of the difficulty of the items.[55] Examiners constructing tests for special purposes thus can control somewhat the distribution of scores on the final test.

In constructing attitude scales, Thurstone has assumed that the scale values assigned to a set of attitude statements are independent of the attitude of the raters. Ferguson, in an experimental investigation, has shown that the median correlation is 0.99 between the weights assigned to a set of items by three groups of judges, quite disparate in their attitudes.[56] It would seem, therefore, that Thurstone's assumption of independence is quite justified.

---

[49] M. A. Saffir, A comparative study of scales constructed by three psychophysical methods, *ibid.*, 2, 1937, 179-198.

[50] P. Horst, Measuring complex attitudes, *J. Soc. Psychol.*, 6, 1935, 369-374.

[51] W. W. Merrill, Jr., Sampling theory in item analysis, *Psychometrika*, 2, 1937, 215-223.

[52] R. F. Sletto, *Construction of Personality Scales by the Criterion of Internal Consistency*, 1937, 1-92.

[53] M. W. Richardson, The relation between the difficulty and the differential validity of a test, *Psychometrika*, 1, 1936, 33-49.

[54] H. A. Voss, An experimental investigation of the relationship between difficulty and reliability in tests. Unpublished Master's Thesis, Fordham University, 1937.

[55] D. A. Walker, Answer-pattern and score-scatter in tests and examinations, *Brit. J. Psychol.*, 26, 1936, 301-308.

[56] L. W. Ferguson, The influence of individual attitudes on construction of an attitude scale, *J. Soc. Psychol.*, 6, 1935, 115-117.

Swineford points out that in determining item validities there is little to choose between biserial r and Pearson r, since the correlation between two sets of coefficients for the same items is high, 0.97, when extremely easy and difficult items have been eliminated, that is, the upper and lower five percentage.[57] She recommends biserial r because of computational ease.

Two papers dealing with sources of errors in examinations and the possibilities of their correction are those of Sandon[58] and Paterson and Raskin.[59] While neither of these are statistical, consideration of the points made will eliminate a great deal of statistical manipulation, so that it seems reasonable to call attention to them here.

The reliability of personality inventories has been studied by Frank,[60] Neprash,[61] and Benton and Stone.[62] Frank has found items marked positively showed less change than those on which the individual was undecided, while Neprash concludes that personality items are unreliable. It appears from the work of Benton and Stone that a period of seven days is sufficient to eliminate or hold constant the effect of the first administration of an attitude scale on a later administration.

*Typology.* A major contribution to the problem of locating types of individuals has been made by Stephenson[63] in a series of excellent papers which are based on a fundamental paper of Thomson's.[64] The basic notion of Stephenson is to correlate individuals rather than tests. This is possible if the number of tests employed is fairly large, that is, twenty or more. Consider a matrix of observations of M columns (tests) and N rows (individuals). Then, normalize each column, that is, express the score in terms of standard scores with a mean of zero and unit standard deviation. It is now possible to take any two individuals, say the i'th and j'th, and treat them as the tests, and find the correlation between the paired, normalized test scores, where the tests are treated as individuals, of which there are M. If the individuals are not alike, the resulting correlation will be zero within sampling errors, where the size of the sample is the number of tests. Obviously, if the number of individuals is large, it is impossible to compute all the intercorrelations, so samples are drawn and types determined in terms of those individuals whose correlations seem to be most alike. The method is extremely simple and seems to be the best method yet proposed for

---

[57] F. Swineford, Biserial r versus Pearson r as measures of test-item validity, *J. Educ. Psychol.*, 27, 1936, 471-472.

[58] F. Sandon, The necessary imperfections of an examination, *Brit. J. Educ. Psychol.*, 5, 1935, 180-193.

[59] D. G. Paterson, E. Raskin, and G. G. Schneidler. Revising test items to increase validity, *J. Psychol.*, 2, 1936, 221-228.

[60] B. Frank, Stability of questionnaire response, *J. Abn. & Soc. Psychol.*, 30, 1935, 320-324.

[61] J. A. Neprash, The reliability of questions in the Thurstone personality schedule, *J. Soc. Psychol.*, 7, 1936, 239-244.

[62] A. L. Benton and I. R. Stone, Consistency of response to personality inventory items as a function of length of interval between test and retest, *ibid.*, 8, 1937, 143-146.

[63] W. Stephenson, Introduction to inverted factor analysis, with some applications to studies of Orexis, *J. Educ. Psychol.*, 27, 1936, 353-367; A new application of correlation to averages, *Brit. J. Educ. Psychol.*, 6, 1936, 43-57; The inverted factor technique, *Brit. J. Psychol.*, 26, 1936, 344-361.

[64] G. H. Thomson, On complete families of correlation coefficients, and their tendency to zero tetrad-differences: Including a statement of the sampling theory of abilities, *ibid.*, 26, 1935, 63-92.

locating types. It offers much of promise to workers in fields such as guidance, attitudes, and interests, and should prove an extremely valuable tool.

*Factor analysis.* The psychologist with only a modicum of statistical training is dazed by the number of factor methods proposed and the vigor with which the followers defend their chosen leader's method. It is a relief to find a competent writer, unbiased by having developed a method of his own, critically examining the various methods. Kellogg[65] has shown that Hotelling's[66] method of obtaining the principal components gives the same results as Thurstone's method of rotation.[67] In a later study, Kellogg[68] examines Kelley's[69] objection to the use of communalities and offers a logical argument for using them. He substantiates this with Flanagan's data[70] by means of a comparison between the methods. Girshick,[71] in an independent investigation, substantiates Kellogg's contention as to the use of communalities. Both of these writers point out that the Method of Principal Axes is the more efficient. The important point in the linking up of these two methods is the fact that Hotelling's work on sampling may now be introduced in part into Thurstone's methods.

Another important comparative study is that of Holzinger and Harman[72] who show the relationships between Holzinger's bi-factor method and Thurstone's method. They also propose a set a standards for judging various factorial analyses which are (a) parsimony of factors, (b) geometric fit, (c) simplicity of structure, (d) independence of factors, (e) statistical fit, and (f) ease of psychological interpretation.

One of the major gaps in factor analysis theory has been the lack of definite criteria for determining the minimum number of vectors necessary to describe a set of inter-correlations. Usually this is done either by inspection or in terms of the standard errors of the residual correlations. Hoel has derived a test which is general in scope and provides a method for determining the minimum number of vectors required.[73] The importance of this paper needs no comment to those using factor methods.

Burt,[74] using a method based on higher moments of variances and covariances, has developed a technique of factor analysis which is faster to use than the methods of Hotelling, Thurstone, and Kelley, since it eliminates the successive approximations inherent in their techniques. This should be welcome information to those working with a large number of variables.

The English workers, while thoroughly cognizant of the methods developed

[65] C. E. Kellogg, The problem of principal components: Derivation of Hotelling's method from Thurstone's, *J. Educ. Psychol.*, 27, 1936, 512-520.

[66] H. Hotelling, Analysis of a complex of statistical variables into principal components, *ibid.*, 24, 1933, 417-441; and 499-520.

[67] L. L. Thurstone, *Vectors of the Mind*, 1935, 1-266.

[68] C. E. Kellogg, The problem of principal components: II. The argument for communalities, *J. Educ. Psychol.*, 27, 1936, 581-590.

[69] T. L. Kelley, *Essential Traits of Mental Life*, 1935, 1-145.

[70] J. C. Flanagan, *Factor Analysis in the Study of Personality*, 1935, 1-103.

[71] M. A. Girshick, Principal components, *J. Amer. Stat. Assoc.*, 31, 1936, 519-528.

[72] K. J. Holzinger, and H. Harman, Relationships between factors obtained from certain analyses, *J. Educ. Psychol.*, 28, 1937, 321-345.

[73] P. G. Hoel, A significance test for component analysis, *Ann. Math. Stat.*, 8, 1938, 149-158.

[74] C. Burt, Methods of factor analysis with and without successive approximation, *Brit. J. Educ. Psychol.*, 7, 1937, 172-195.

in America, are still devoting a great deal of energy to the tetrad method and to the question of 'g' and 's,' as shown by recent papers of Thomson,[75] Stephenson,[76] and Piaggio.[77] Spearman[78] has recently examined the question, why should a variable be expressed as a sum rather than as a product of two variables? He concludes that, for the case examined, either assumption leads to precisely the same propositions.

Psychologists, even as you and I, are prone to hail each new development in statistics as a magic touchstone to solve their difficulties. When the method of correlation became available, there was a rush to correlate every conceivable measure with every other measure. Today the mode is factor analysis, and it is only right that this new tool should be used by psychologists in an attempt to solve the varied and intricate problems with which they are confronted. A brief summary of some of the many applications of factor analysis will give an idea of the range of problems to which it has been applied. Dunlap attempted to apply it to the problem of the organization of learning in chickens,[79] while Thorndike used the method in studying the organization of behavior in albino rats.[80] Kulp and Davidson have used the method in studying international attitudes,[81] and Whisler has attempted to locate generalized attitudes by this technique.[82] Guilford and Guilford,[83] in a study of personality, found five independent factors; whereas, in a more recent study, Mosier describes eight factors.[84] Student's traits have been studied by Rexroad with considerable success.[85] This study is essentially one of typology like those of Stephenson who also uses factor analysis techniques.[86] The method has been used in learning problems by Perl, who identified two factors in the learning situation.[87] Not even the nursery school has escaped the application of the method. Koch,[88] studying measures of activeness for nursery school children

---

[75] G. H. Thomson, On measuring g and s by tests which break the g-hierarchy, *Brit. J. Psychol.*, 25, 1934, 204-210.

[76] W. Stephenson, On Thomson's theorem for measuring g by overlapping tests, *ibid.*, 25, 1935, 490-494.

[77] H. T. H. Piaggio, Approximate general and specific factors without indeterminate parts, *ibid.*, 25, 1935, 485-489.

[78] C. Spearman, Abilities as sums of factors, or as their products, *J. Educ. Psychol.*, 28, 1937, 629-631.

[79] J. W. Dunlap, The organization of learning and other traits in chickens, *Comp. Psychol. Monog.*, 9, 1933, (no. 4), 1-55.

[80] R. L. Thorndike, Organizations of behavior in the albino rat, *Genet. Psychol. Monog.*, 17, 1935, 1-70.

[81] D. H. Kulp, and H. H. Davidson, The application of the Spearman two-factor theory to social attitudes, *J. Abn. & Soc. Psychol.*, 29, 1934, 269-275.

[82] L. D. Wisler, Multiple-factor analysis of generalized attitudes, *J. Soc. Psychol.*, 5, 1934, 283-297.

[83] J. P. Guilford and R. B. Guilford, Personality factors S, E, and M, and their measurement, *J. Psychol.*, 2, 1936, 109-127.

[84] C. I. Mosier, A factor analysis of certain neurotic symptoms, *Psychometrika*, 2, 1937, 263-288.

[85] C. N. Rexroad, A factor analysis of student traits, *J. Educ. Psychol.*, 28, 1937, 153-156.

[86] Stephenson, *op. cit.*

[87] R. E. Perl, An application of Thurstone's method of factor analysis to practice series, *J. Gen. Psychol.*, 11, 1934, 209-212.

[88] H. L. Koch, A multiple-factor analysis of certain measures of activeness in nursery school children, *J. Genet. Psychol.*, 45, 1934, 482-487.

used a multiple factor analysis, and in a study of children's annoyances, the method is applied by Carter, Conrad, and Jones.[89]

In concluding this summary of recent work in factor analysis, a word of caution is advisable. Thomson,[90] in the conclusions of a study on vocational guidance in which the method was used, says that "the preceding sections make it quite clear that when vocational guidance proceeds by giving the candidate a number of tests which have previously been given to persons already engaged in the occupation, the use of factors has no mathematical justification whatever." So many studies have appeared in which factorial methods are misused that Thurstone,[91] feeling responsible for his part in the development of the techniques, has seen fit to point out a number of conditions which often are neglected. The conditions that he considers neglected most often are: tests used in factor analysis contain too many factors, that is, general tests should not be used; failure to rotate the axis into a simple configuration, which is necessary if the test loadings are to remain invariant, and psychological meaning is to be attached to the factors; and, finally, the use of too few tests containing each factor.

Thurstone has stated that, if the test vectors are rotated into coördinate planes, the resulting loadings remain invariant as a test is moved from one test battery to another.[92] This leads him to believe that the unitary abilities have genetic significance, and that it will be possible eventually to identify the special abilities genetically. Only one study, that of Smart,[93] has come to the writer's attention in which a test was systematically shifted from one battery to another. Smart found that the test loadings varied as the test was shifted. Unfortunately, he did not rotate the vectors into a simple configuration, so that the question still awaits experimental verification.

Notwithstanding these warnings, there can be no question that the method of factor analysis provides an unusually powerful means of attacking highly complex problems, and it should be used whenever the conditions are appropriate.

Two other papers should be called to the reader's attention, although they fall into none of the above classifications. Tilton has developed a measure of overlapping which gives the worker a more accurate picture than heretofore available, and which has a wide variety of applications in psychological and educational work.[94] An unusual study of growth is that of Burgess who has made a synthetical statistical analysis of sixty published reports on height.[95] Workers interested in the problems of

---

[89] H. D. Carter, H. S. Conrad, and M. C. Jones, A multiple factor analysis of children's annoyances, *ibid.*, 47, 1935, 282-298.

[90] G. H. Thomson, Some points of mathematical technique in the factorial analysis of ability, *J. Educ. Psychol.*, 27, 1936, 37-54.

[91] L. L. Thurstone, Current misuse of the factorial methods, *Psychometrika*, 2, 1937, 73-76.

[92] L. L. Thurstone, Unitary abilities, *J. Gen. Psychol.*, 11, 1934, 126-132.

[93] R. C. Smart, The variation in pattern of factor loadings, *J. Educ. Psychol.*, 28, 1937, 55-64.

[94] J. W. Tilton, The measurement of overlapping, *J. Educ. Psychol.*, 28, 1937, 656-662.

[95] M. A. Burgess, The construction of two height charts, *J. Amer. Stat. Assoc.*, 32, 1937, 290-310.

growth, such as of intelligence or learning, should examine the method employed in this study.

Occasionally one finds in the current literature the use of the wrong formula. In a recent published study between correlated variables, although correlations were reported, no use was made of them in computing the significance of the differences. While this type of error is serious in many cases, it is negligible as compared with the error of applying statistical method to non-statistical data. Jellinek,[96] in a paper designed for psychiatrists, ably illustrates this type of error. Often the error lies in the failure of the worker to differentiate between necessary and sufficient conditions, with a result that undue emphasis is given the necessary conditions.

Certain significant trends in statistical methodology seem apparent from the examination of the studies reported here, and of others from which these are selected. They are the increasing use of the method of rank correlation where normality cannot be assumed, or where certain hypotheses as to order are being examined; the slow but steady increase of the use of analysis of variance; and the sudden but great increase in the use of chi square. Analysis of variance should be used more often in psychological problems, for it is a simple and very powerful tool. The importance of chi square has been recognized by biologists as evidenced by Snedecor's paper.[97] Fisher states that chi square "is the great contribution to statistical methods by which the unsurpassed energy of Professor Pearson's work will be remembered."[98] Pearl expresses the same opinion when he says that Pearson's important technical contributions fall into four categories, and that one of these is the development of the chi square test.[99]

In conclusion, I can only extend my sympathy to the psychologist of the future, for it seems as if he must first be a mathematician, then a statistician, then a physiologist, then a physicist, and, if he is not dead of old age by then, a psychologist.

University of Rochester JACK W. DUNLAP

## TWO EQUATIONS FOR THE STUDY OF VARIABILITY

It is frequently desirable to compute the sum or the average of all the differences possible in a distribution of scores. An equation for obtaining the sum or the average of all such differences may be derived in the following manner.

The scores in a distribution may be arranged in their order of magnitude from the highest score to the lowest. These may be represented by the following series:

$$k_n,\ k_{n-1},\ k_{n-2},\ \ldots\ k_1 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots [1]$$

where $k_n$ is the largest score, $k_{n-1}$ is the next largest, and so on. The sum of all the possible differences among these numbers may be represented as

$$D = (k_n - k_{n-1}) + (k_n - k_{n-2}) + \ldots + \cdot\ (k_{n-1} - k_{n-2}) + \ldots + \cdot\ (k_2 - k_1) \quad [2]$$

It should be noted that $k_n$ will appear $(n-1)$ times and will always be positive;

[96] E. M. Jellinek, Some uses and abuses of statistical method in psychiatry, *Biom. Bull.*, 1, 1937, 97-108.

[97] G. W. Snedecor, The improvement of statistical techniques in biology, *J. Amer. Stat. Assoc.*, 31, 1936, 690-701.

[98] R. A. Fisher, *Statistical Methods for Research Workers*, 5th ed., 1934, 23.

[99] R. Pearl, Karl Pearson: 1857-1936, *J. Amer. Stat. Assoc.*, 31, 1936, 653-664.

$k_{n-1}$ will have positive values in (n—2) instances and negative values in 1 instance; $k_{n-2}$ will be positive (n—3) times and negative 2 times, etc. Therefore, by rearranging the sums and differences we obtain

$$D = (n-1)k_n + (n-3)k_{n-1} + (n-5)k_{n-2} + \ldots + \cdot\ (1-n)k_1 \quad \ldots\ldots\ldots\ldots [3]$$

In order to obtain the average difference we need only to divide the equation by $n(n-1)/2$. The latter value will give us the number of differences that exist.

It should be noted that equation [3] is symmetrical so that $(n-1) = -(1-n)$; $(n-3) = -(3-n)$; etc. The equation may therefore be simplified to

$$D = (n-1)\ (k_n - k_1) + (n-3)\ (k_{n-1} - k_2) + \ldots\ldots\ldots\ldots [4]$$

These combinations may be made in all cases since in a series with an odd number of terms the middle term will be zero. In equation [4] the last term will be $2(k_{n+2}/2 - k_{n-1}/2)$ if the series is odd and $(k_{n+2}/2 - (k_n/2)$ if the series is even.

Substituting in the formula is a simple matter as is shown by the following example. In an experiment which required the subjects to rank a series of occupations for five different characteristics, a given occupation received the following ranks: 2, 3, 4, 6, and 9. To obtain the sum of all the differences between pairs of ranks we find by substituting in [4]: $D = 4(9-2) + 2(6-3) = 28 + 6 = 34$. Since there are five scores, the number of differences, $n(n-1)/2 = 10$. The average is therefore 3.4. If two or more numbers of equal magnitude are present in the distribution, the formula is applicable provided that ranks are not split. Thus if our series were 2, 2, 2, 6, and 9 we could assign the third position to one of the 2's, the second to another, and the first (*i.e.* the lowest position) to the third. D would then be equal to $4(9-2) + 2(6-2) = 28 + 8 = 36$.

Should the investigator wish to have the sum of the *squares* of all possible differences, thus weighting the more extreme differences, a formula giving the desired measure may be derived in the following manner:

It is possible to express the standard deviation of a distribution by utilizing deviations from an arbitrary mean. For the square of the standard deviation we have

$$\sigma^2 = \Sigma(X-M')^2/n - c^2 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots [5]$$

where $c^2$ represents the square of the difference between the obtained and the arbitrary mean. This formula may be applied directly to the problem at hand. If we substitute one of the obtained scores for M', for example, $X_1$, we may by transposing obtain an expression for the sum of the squares of the differences from the number substituted, *i.e.* $X_1$. If now we successively substitute $X_2$, $X_3$, etc., we will end up with a formula expressing twice the sum of the squares of all the differences. From this equation we may get a number of different but equivalent expressions. Substituting any one of the obtained scores we get

$$\sigma^2 = \Sigma(X-X_1)^2/n - (M-X_1)^2 \quad \ldots\ldots\ldots\ldots\ldots\ldots [6]$$

where $X_1$ is any obtained measure, M is the obtained mean, and $(M-X_1)$ represents the difference between the obtained and the arbitrary mean. Solving for $\Sigma(X-X_1)^2$ we have

$$\Sigma(X-X_1)^2 = n\sigma^2 + n(M-X_1)^2 \quad \ldots\ldots\ldots\ldots\ldots\ldots [7]$$

$$\Sigma(X-X_1)^2 = n\sigma^2 + nM^2 - 2nMX_1 + nX_1^2 \quad \ldots\ldots\ldots\ldots [8]$$

Equation [8] expresses the sum of the squares of the differences of the obtained

measures from any one of the obtained measures. By summating we shall obtain a function of all the differences.

$$\Sigma[\Sigma(X-X_1)^2] = S = n^2\sigma^2 + n^2M^2 - 2nM\Sigma X + n\Sigma X^2 \quad [9]$$

It should be noted that $X_1$ and $X_1^2$ when summated simply become $\Sigma X$ and $\Sigma X^2$ respectively. The remaining terms are repeated n times each. Since $M = \Sigma X/n$, $n^2M^2$ reduces to $(\Sigma X)^2$ while $-2nM\Sigma X$ reduces to $-2\ \Sigma X)^2$. Therefore:

$$S = n^2\sigma^2 - (\Sigma X)^2 + n\Sigma X^2 \quad [10]$$

However, $n\Sigma X^2 - (\Sigma X)^2$ is another expression for $n^2\sigma^2$ since with the arbitrary mean equal to zero we obtain:

$$\sigma^2 = \Sigma X^2/n - M^2 = \Sigma X^2/n - (\Sigma X)^2/n^2 = [n\Sigma X^2 - (\Sigma X)^2]/n^2 \quad [11]$$

Therefore $S = 2n^2\sigma^2$. [12]

S is twice the desired sum since each difference squared is obtained twice. Each X is in turn used as a base from which the other X's are subtracted, and when not used as such a base is subtracted from the remaining X's. Therefore each difference between a pair of X's occurs twice. In n instances, a number has been subtracted from itself, reducing the differences to those in which we are interested, namely, $n(n-1)$. If the actual sum is represented by S', we have $S' = n^2\sigma^2$. [13] The average of the sum of the squares of all differences will then equal:

$$2n\sigma^2/(n-1) \quad [14].$$

Equation [13] may be expressed in other ways, depending on the convenience of the investigator. It may be given as $S' = n\Sigma X^2 - (\Sigma X)^2$, [15] or if the mean is available $S' = n\Sigma X^2 - n^2M^2$ [16]

The example used above for equation [4] will be used here to illustrate substitution in equations [13], [15], and [16]. We have five rankings 2, 3, 4, 6, and 9. $\sigma = 2.48$; $\sigma^2 = 6.15$. Therefore, substituting in [13] $S' = 25(6.15) = 153.75$. The average of this sum is 15.38. For equation [15] we find $\Sigma X = 24$; $\Sigma X^2 = 146$. $S' = 5(146) - (24)^2 = 730 - 576 = 154$. The average of the squares of the differences is 15.4. For equation [16], $M = 4.8$ and $M^2 = 23.04$. $S' = 5(146) - 25(23.04) = 730 - 576 = 154$. The average is 15.4. The discrepancy between the results obtained from using equation [13] as compared with [15] and [16] can be narrowed by using a greater number of decimal places.

Equations [4], [13], [15], and [16] are most readily applied in problems dealing with individual variability. With equation [4], a good estimate of variability may be obtained for a small series of measures where it is important to compare each measure with every other one.

The other equations may be used for the same purpose, although ordinarily it is not desirable to weight the more extreme differences. Which of the equations is to be used may often depend on other aspects of the statistical treatment. If it is not necessary to obtain the standard deviation, equations [15] and [16] are most convenient. If the mean must be obtained, equation [16] may be used provided that the standard deviation is not necessary.

It should be pointed out in connection with [14] that a close approximation to the average of the squares of the differences may be gotten when the number of cases is fairly large by allowing n to be equal to (n—1). Then the expression reduces to $2\sigma^2$. [17]

Of all the equations [4] is probably the most useful. The others will represent

no different order of variability among a group of individuals than that represented by individual standard deviations. Of course, if absolute magnitudes are important they are an improvement on the standard deviation. Equation [4], however, is easily utilized and is not limited by the errors that will usually be incurred by introducing functions of the standard deviation in small samples.

College of the City of New York MAX HERTZMAN

## AN OPTICAL ILLUSION OF PERSONAL MAGNETISM

An elementary laboratory student of psychology, Mr. Wei-Dun Ho, discovered accidentally three years ago, while dressing before the mirror one morning, a peculiar visual phenomenon, which he himself could not then very well explain. One afternoon when we met in the discussion-class, he demonstrated this phenomenon to us and asked the writer for a psychological explanation. At first we thought that it must be one of the geometrical optical illusions, but after some reflection the writer was convinced that it is truly physical in nature, and hence names it now as an "optical illusion of personal magnetism."

The best way to obtain this phenomenon is to use a medium-sized mirror, 15 in. × 20 in. and ¼ in. thick, on the surface of which a thin, even layer of chalk dust is sprinkled. If one looks closely into this mirror, at the same time paying attention to the pheriphery near the fixation point, one will notice that one's own penetrating eye-sight seems to be endowed with magnetic induction, causing the dust particles surrounding the fixation-point to radiate in lines from that point.

The association between this visual illusory effect and magnetic lines of force is spontaneous and naïve. Many subjects are even bewildered to find that it is independent of eye movement. Wherever the eyes wander together with the head, that illusory radiation of magnetic lines of force follows their fixation-point. The writer demonstrated later this phenomenon to several psychologists and physicists, besides many elementary and advanced psychology students. Most of the observers got the effect readily, but none of them offered offhand the correct explanation. Those who had any explanation at all explained it in terms of either purely subjective or entirely irrelevant physical factors. Some actually believed that it is due to personal magnetism of one's own eyesight. According to them, our eye has a subtle emanation that pervades space, a magnetic fluid that radiates from one's own eyesight and may influence the minds and bodies of others or even physical objects.[1] In fact, this explanation seems to be so real that, while observing the phenomenon, one cannot think otherwise.

In reality, this optical effect differs fundamentally, however, from ordinary illusion. For in ordinary illusion, there is nothing actually wrong with the stimuli. Here, however, the stimuli are physically half real and half unreal. It is entirely a physical affair and is sense-perception pure and simple. The true explanation may be given with the aid of Fig. 1.

As the eye fixates on particle *A* on the surface of the mirror, its projected image *a*, being directly underneath, will naturally be reflected to the same place as *A*,

[1] Cf. J. E. Coover, The feeling of being stared at, *Experiments in Psychical Research*, 1917, 144-167.

*i.e.* the reflected image *A'* will coincide with *A*. In peripheral vision, however, the projected image *b* of another particle *B* nearby will be reflected on the surface as *B'*, midway between *A* and *B*. Hence one sees on the surface of the mirror three particles in line, two real ones, *A* and *B*, and a third reflected particle *B'*, the image of the projection *b* of particle *B*, lying between *A* and *B*. Similarly, the image *C'* of the projection *c* of another particle *C* will lie between *B* and *C;* the image *D'* of the projection *d* of still another particle *D*, between *C* and *D*, and so on and so forth. When we come to particle *G*, however, the image *G'* of its projection *g*

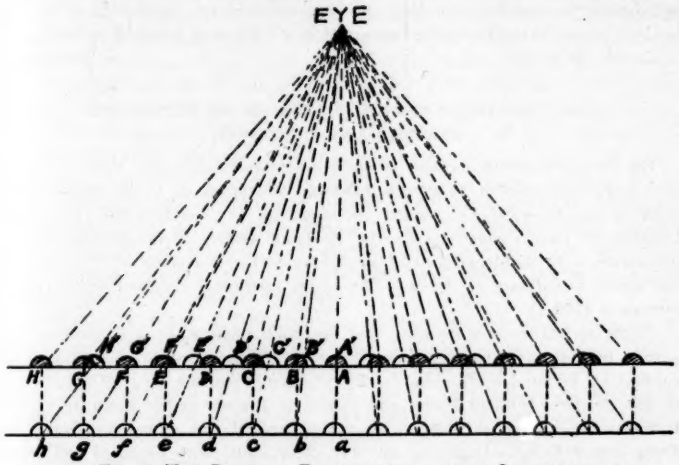


FIG. 1. THE PHYSICAL EXPLANATION OF THE ILLUSION
*A, B, C, D,* etc., dust particles on the surface of the mirror; *a, b, c, d,* etc., images on the silvered back of the mirror; A', B', C', etc., images projected to the surface of the mirror.

will exactly coincide with the particle *F* on the surface, and the image *H'* of the projection *h* of the last particle *H*, instead of being between *G* and *H*, will lie between *F* and *G*. Thus *A–B'–B–C'–C–D'–D–E'–E–F'G'–F–H'–G–H* taken altogether, naturally give the impression of a continuous line radiating in one direction from particle *A*, the fixation-point.

In the same fashion, a second line will radiate from the same center *A* in another direction, a third one in a third direction, and so on, with as many lines radiating from the same center *A* as there are particles immediately around it.

National Tsing Hua University SIEGEN K. CHOU

## THE 'MAGNETIC EYE' ILLUSION

Experimental variation of the 'magnetic eye' illusion—to give the illusion reported above a simple name—confirms Professor Chou's explanation.

*(1) Negative test.* The illusion fails to appear when the chalk particles are

applied to an opaque surface, such as black cardboard. The doubling of the particles by the mirror is essential to the illusion.

*(2) India ink.* The chalk is replaced with India ink, sprinkled by scraping the pen against a brush. Again, the illusory radii appear on a mirror-surface and fail to appear on white cardboard.

The explanation from physical optics requires a psychological supplement. Wertheimer has discussed the rôle of spacing in the perception of objects.[1] His *Faktor der Nähe* is operative in the present illusion, in which perception of linearity is favored by the closer spacing due to the mirror-images. The same factor of spacing, distance or proximity has long ago been noted in the perceptions of duality: in visual acuity, in the two-point limen, and in the limen of temporal succession.

Cornell University S. FELDMAN

## THE THIRTEENTH ANNUAL MEETING OF THE MIDWESTERN PSYCHOLOGICAL ASSOCIATION

The thirteenth annual meeting of the Midwestern Psychological Association was held at the University of Wisconsin, Madison, Wisconsin, on Friday and Saturday, April 22 and 23, 1938. The register was signed by 199 members and 215 visitors, bringing the known attendance to 414. A special feature of the meeting was the celebration, with the department of psychology at the University of Wisconsin, of the fiftieth anniversary of the founding of its psychological laboratory by Joseph Jastrow in 1888.

At the annual business meeting, held on Friday evening, 48 new members were elected, bringing the present membership to 480. Announcement was made of the election of Edmund S. Conklin to the presidency for the year 1938-1939, and of the election of Elmer Culler and Sidney L. Pressey to the executive council. It was voted to publish abstracts of papers read at the annual meeting in the *Psychological Bulletin* beginning in 1939. Resolutions were proposed and passed by a vote of 31 to 24 protesting the holding of the Twelfth International Congress of Psychology in Vienna in 1941. It was announced that the next annual meeting of the Association will be held at the University of Nebraska, Lincoln, Nebraska, May 5 and 6, 1939.

Twelve sessions were held at which 64 scheduled papers were read, a great variety of fields being covered: social behavior, sensation and perception, affectivity and motivation, electrophysiology, memory, personality, mental development, educational psychology, conditioned response, applied psychology, animal psychology, and speech. Nine formal symposia were held with invited papers on the topics: operationism, learning, extra-sensory perception, principles of development, first course in psychology, climate and human behavior, factor analysis, personality adjustment, family relations, and contemporary social problems. On Saturday afternoon the customary round-table was held at which "reports from laboratories" were received.

On Friday evening a Golden Anniversary dinner was held at the Park Hotel honoring Joseph Jastrow, Professor Emeritus, and Clark L. Hull, formerly of

[1] M. Wertheimer, Untersuchungen zur Lehre von der Gestalt, *Psychol. Forsch.*, 4, 1923, 301-350.

Wisconsin. Jastrow's address following the dinner, entitled "Stages in the progress of psychology," was enjoyed for its reminiscences of historical events and its mature judgments of the direction that psychology of the present is taking. Hull's address on "Adventures in psychological theorizing" related his attempts to rationalize the problem of memorizing.

Following the annual dinner on Saturday evening, at which Harvey A. Carr presided, the presidential address was given by Arthur G. Bills. The title of his address was "Changing views of psychology as science." In substance he said, quoting from an abstract supplied by Dr. Bills, that "changes are taking place in the views of psychologists as to what are the basic criteria of a scientific psychology. A suggested minimum platform for psychology as science would include the following criteria: psychology as a science is mechanistic, analytic, quantitative, nomothetic, empirical, and uses operational concepts. But 'mechanistic' no longer implies 'machinelike'; 'analytic' no longer requires the discovery of real concrete entities nor irreducibles; and 'quantitative' refers to the finished stages of the formulation of laws. The nomothetic criterion precludes a purely individual or personalistic psychology. The operational criterion, for which too much is often claimed, merely insures the experimental derivation of concepts, without specifying the nature of the experimental operations, further than that they must be fundamental for the particular science. The empirical criterion, the most important of all, heretofore taken for granted, is threatened by a rising counter-revolution which may annihilate it, unless scientists of empirical ideals come to the rescue."

All sessions were well attended and an unusual seriousness of purpose and a generous participation in the discussions was evident. To the local committee, composed of R. W. Husband, H. F. Harlow, and Walter F. Grether, to V. A. C. Henmon and all members of his department, and to the University of Wisconsin were extended the sincere thanks of the Association for the splendid hospitality and arrangements which made the meeting enjoyable and successful.

University of Nebraska J. P. GUILFORD

## THE NINTH ANNUAL MEETING OF THE EASTERN PSYCHOLOGICAL ASSOCIATION

The ninth annual meeting of the Eastern Branch of the American Psychological Association was held in New York City on April 1 and 2, 1938. A two-day session was held for the first time. On Friday the meetings and banquet, at which the president's address was given, were held at the Hotel Commodore; and on Saturday the meetings were held on the upper campus of New York University in University Heights. The total registration numbered 620: 343 members and associates, and 277 guests, which is the largest number ever to attend these meetings.

The program, which is to be printed in the *Psychological Bulletin*, consisted of 107 papers. These papers were presented in 15 sectional meetings. Six sections held simultaneous meetings Friday afternoon at the Hotel Commodore at which 47 papers were read. Following these meetings there was a special section for the showing of instructional films. Seven simultaneous sections met Saturday morning at which 55 papers were read. Saturday afternoon, after the annual business meeting, a general session was held at which the following papers were read: Carney Landis, "Methods, techniques, and conclusions in contemporary abnormal psychology;" Jack Dunlap, "Recent statistical advances and their applications in

psychology;" George Kreezer, "Present status of research on brain potentials."

At the annual business meeting, the following recommendations of the Board of Directors were approved: (1) The name of the organization was changed from the Eastern Branch to "The Eastern Psychological Association," with the subtitle "Affiliated with the American Psychological Association." (2) A two-day session was ordered for 1939. (3) The invitation for the meeting on March 31 and April 1, 1939 from Bryn Mawr College was accepted. (4) The Association will discontinue the practice of printing abstracts of scientific papers in the *Psychological Bulletin*.

The following officers and committee members were elected. President (1938-39), Karl M. Dallenbach, Cornell University; Directors (1938-41), Harry W. Helson, Bryn Mawr College; Charles W. Bray, Princeton University; (1938-40), Kurt Koffka, Smith College. Program Committee, T. W. Forbes, Harvard University; E. Heidbreder, Wellesley College; C. Jacobsen, Cornell Medical College. Nominating Committee, I. Lorge, Teachers College, Columbia; G. Kreezer, Vineland Training School; Myrtle B. McGraw, Babies Hospital, Columbia.

The area in which notices for the meeting were sent this year extended from the New England States and Canada as far south as Washington, and west to Buffalo and Pittsburgh. Approximately 1140 eligible members reside in this territory.

The Association is indebted to the members of the staff at New York University, University Heights, for the excellent arrangements for this two-day meeting.

University of Pennsylvania H. S. OBERLY

### CHAPEL HILL MEETING OF THE SOCIETY OF EXPERIMENTAL PSYCHOLOGISTS

The thirty-fourth meeting of the Society of Experimental Psychologists was held on April 7 and 8 at the University of North Carolina, Chapel Hill, N.C., under the Chairmanship of Professor John F. Dashiell. It was attended by 23 members from 17 institutions.

After a short business session, informal reports of work in progress were received from the following institutions and men: Brown (Graham and Schlosberg), California in Los Angeles (Dunlap), Columbia (Landis), Cornell Medical (Jacobsen), Harvard (Beebe-Center, Boring), Illinois (Culler), National Research Council (Cobb, Johnson), North Carolina (Dashiell), Pennsylvania (Fernberger), Princeton (Bray, Langfeld, Wever), Rochester (Carmichael), Rutgers (Pratt), Swarthmore (Köhler), Vanderbilt (Lanier), Washington (Nafe), Wesleyan (McGeoch), and Yale (Marquis, Miles).

At a dinner given to the Society by the University of North Carolina, the Howard Crosby Warren Medal for outstanding research in experimental psychology was presented to Professor Elmer A. Culler of the University of Illinois "for his work on the fundamental mechanisms of hearing and on the physiological basis of the conditioned reflex."

The next annual meeting of the Society is to be held at Princeton University on Monday and Tuesday, April 3–4, 1939. Professor Herbert S. Langfeld of Princeton University was elected chairman of the Society for the coming year and Professor Samuel W. Fernberger of the University of Pennsylvania was reëlected Secretary and Treasurer for a three-year period. S. W. F.

## Professor Bentley Retires from Teaching

Professor Madison Bentley, an editor of this JOURNAL since 1903, became emeritus professor of psychology of Cornell University on June 30, 1938. On the occasion of his retirement, his students and colleagues of the universities of Illinois and Cornell presented to him, at a dinner on June 11, 1938, an oil portrait of himself painted by Olaf M. Brauner (reproduced in the frontispiece of this number) and a book of testimonial letters.

The first fourteen and the last ten of Professor Bentley's forty years of service to psychology were passed at Cornell; the intervening sixteen years, at the University of Illinois. Following his doctorate at Cornell in 1898, he served as instructor, assistant professor (1902), and head of the undergraduate division of the department (1909). In 1912, he went to the University of Illinois as professor and head of the department of psychology. Except for the period of the World War, when he was captain in the Aviation Corps of the U. S. Army, he remained at Illinois until 1928. In that year he was called back to Cornell as Sage Professor of Psychology and chairman of the department in succession to Professor E. B. Titchener.

In addition to his editorship of this JOURNAL (coöperating editor from 1903 to 1925 and co-editor since 1926), Professor Bentley was also editor of the *Psychological Index* (1916-1925) and the *Journal of Experimental Psychology* (1926-1929) and associate editor of the *Journal of Comparative Psychology* (1920-1935).

An honorary LL.D. degree was conferred on Professor Bentley by the University of Nebraska in 1935. He was president of the American Psychological Association in 1925, vice-president of the section on psychology of the American Association for the Advancement of Science in 1926, and chairman of the Society of Experimental Psychologists in 1929. He is also a fellow of the National Institute of Psychology.

Professor Bentley's service to psychology will not end with his retirement from teaching. The scene of his activities is merely shifted to Washington, D.C., where he will act as consultant for psychology on the staff of the Library of Congress.

K. M. D.

## John Edgar Coover: 1872 – 1938

John Edgar Coover, Emeritus Professor of Psychology at Stanford University since June, 1937, died at his home on the Stanford campus on February 19, 1938. Dr. Coover's academic career was almost wholly confined to Stanford University, and it was from that institution that he received his professional training.

He was born near Remington, Indiana, of German parentage on March 16, 1872. His parental line can be traced back to Dieterich Kober, a native of the German Palatinate, who arrived in Philadelphia in 1730.

When twenty-two years old, Coover satisfied an ambition which had been maturing for some time when he entered the State Normal School at Greeley, Colorado. It was no easy task for a farm boy without financial backing to find the wherewithal to attend college in Colorado in the pre-panic years of 1890-92. His elementary schooling had been meager, and his boyhood and adolescent years appear to have been filled with more than the usual amount of hard work. His own modest characterization of this portion of his life indicates the possession of a single-minded

devotion to an educational ideal that was extraordinary.[1] This period, however, was chiefly notable for the development of three vocational interests which were to play an important rôle in his later professional and scientific career. He became interested in telegraphy, Pitmanic shorthand, and printing, and acquired a craftsman's competence in the use of all three skills. They not only served his financial need both at the Normal School and later at the University, but furnished the the materials and the problems for his later research.

In 1937, after 43 years, this period still seemed to possess or had acquired glamour, but not without a price. Coover wrote:

> Those were great days. The capacity for work and tenacity of purpose they engendered stood me in good stead in working my way through Stanford, the first two years by a bit of stenography and typing and a lot of printing, the next four by telegraphy. (One never knows when he may need skills, or knowledge, once acquired.) These traits seem to have pursued me during my whole life, but at a price: I never had the time to learn the social devices by which gentlemen kill time—dancing, cards, golf, lounging.

The teaching problem and the teaching process fascinated Coover at this time, and the "glorious privilege to assist young minds to awake into fruitful and gratifying knowledge, clear understanding and good judgment," became a dominate enthusiasm and furnished a motivation which persisted throughout life. To serve this purpose the study of psychology seemed essential, but philosophical problems intruded. Coover said:

> My Junior year at the State Normal, however, brought me to grips with an important problem. A brilliant young instructor gave a year's course (five days a week) in psychology. And there was a worm in its core. I read every relevant book in the library, and located the difficulty; it was a disregard of the mind-body problem. The course was an easy, clear mechanistic system borrowed from the French (mostly Ribot). Stoutly against it were the giants of Germany: Wundt, Külpe, and even James in America. The course seemed to be popular; but the best students were wrecked in it, their high ambitions extinguished. The child is a little mechanism, under control of the springs of its nervous system; it grows mechanically by reason of changes of the stimuli in its presence brought by the teacher; but the teacher is, too, a mechanism; the changes he effects in the child's environment are brought about by the mechanized nervous system of his own. No place for character development, only habit formation. This dogma was not only untrue, it was dangerous. I should become a teacher of psychology. But I must have the best preparation: an Ann Arbor A.B., a Harvard M.A., a Leipzig Ph.D.

The stage was thus set and the course plotted. The outward and visible signs of Coover's progress along this self-appointed path may be briefly recorded. He was librarian and secretary to the President of the Greeley Normal School from which he received the Ped.B. in 1898. Following a year as school principal at Como, Colorado, he matriculated at Stanford University. Here he received the A.B. in Philosophy in 1904 and the A.M. in Psychology in 1905. His first teacher

[1] The writer had access to several sources of biographical material in the form of notes and summaries prepared by Coover for various purposes. The chief source of information was a "Chronological Outline" prepared by him about two months before his death. He had organized this material in connection with a request from the *National Cyclopedia of American Biography*. This "outline" is an intimate, informal account, in the first person, and, although incomplete, it contains much revealing material. All quotations in this note are taken from it.

of philosophy was Lovejoy, and summer sessions at the University of California found him in the classes of Royce, Palmer, and J. M. Baldwin.

In 1910, Coover returned to Stanford as a graduate student after five years of teaching as a high-school principal. He was appointed Assistant in Psychology and in 1912 he received the doctorate and became Fellow in Psychical Research.

The founding of the Psychical Research Fellowship at Stanford was an important event, not only in the life of John Edgar Coover, but also in the development of research in a highly controversial field. Thomas Welton Stanford, brother of Senator Stanford, founder of the University, was a devout believer in spiritualism, but was willing to endow a program of research, the outcome of which might prove unfavorable to his cherished beliefs. Ten thousand pounds were placed at the disposal of the University, the income of which was to support the research. The selection of a Fellow was left in the hands of President Jordan and Professor Frank Angell, Chairman of the Department of Psychology. Coover was appointed and held the fellowship continuously until his retirement in 1937.

He proceeded slowly through the levels of the academic hierarchy. In 1914 he was made Assistant Professor, in 1921 Associate Professor, and in 1930 Professor of Psychology. During the World War he was Captain in the U. S. Army Sanitary Corps, and was attached to the Psychological Division of the Medical Research Laboratory at Mineola, Long Island. Here he introduced experimental methods and statistical procedures for determining an aviator's ability to maintain motor coordination and attention under deprivation of oxygen. In 1927 he was visiting Associate Professor at the Johns Hopkins University.

Coover's primary research interests were in the fields of psychophysics and cognition. The closely allied interests in learning and teaching were not sporadic, or the result of chance, but sprung from his early concern with the educational and ethical implications of the mind-body problem, and his almost passionate faith in education and the high task of the teacher. These were integrating factors in Coover's professional life and all his activities as scientist and teacher were related to them.

His scientific projects fall roughly into four classes: psychophysics, learning and cognition, statistical procedures and psychical research. In psychophysics Coover investigated visual thresholds and the influence of secondary criteria on the lower grades of certainty and upon "guesses." His studies of the learning process include a collaborative study with Professor Frank Angell on the general practice effect of special exercise and his doctoral dissertation on formal discipline. His later studies of simple motor performances and of speed of writing symbols in substitution tests had practical, but incidental, expressions in studies of expert typing performance, the production (with E. G. Wiese) of a manual of instruction for typists, and the invention of a system of speed writing and typing published in 1937 by the Stanford University Press under the title of *Note Script and Note Typing*.

Coover's statistical researches were primarily confined to the development of simplified methods for use in the analysis of laboratory data. A book in manuscript form contains the results of this work.

Psychical research, of course, represented for him a major activity, but it is probable that Coover regarded this as a professional necessity rather than a life interest. His *Experiments in Psychical Research*, published in 1917, must be re-

garded as his major scientific contribution, however, and the contemporary flurry of interest in this field is likely to enhance the prestige of this monograph.

Coover's published output was comparatively small. A meticulous, almost perfectionist attitude with reference to his research, together with a certain tendency to withdraw from intellectual and social contacts—in part the result of a feeling of inferiority and in part the result of environmental factors—may have accounted for the paucity of his publication. He anticipated a time when, free from distraction, he would find those circumstances which would favor the complete exposition and interpretation of the results of twenty-five years of research. That some of this mass of accumulated material may be salvaged is the hope of those persons who have had the opportunity to examine it.

Probably no student—certainly no graduate student—who penetrated to the third floor Tower Room at Stanford left without a sense of having made a contact with a rich and aristocratic mind. Coover could not and would not be hurried. In speech and action he was slow but not ponderous, certain but not dogmatic. To those who were his intimates—too few in number!—he was a gay, whimsical and loyal companion.

University of California at Los Angeles FRANKLIN FEARING

# BOOK REVIEWS

*Psychology Down the Ages.* By C. S. SPEARMAN. New York, Macmillan Co., 1937. Pp. xi, 454; vii, 355.

In these two volumes the author gives a summary of the development of psychological theory and uses it as a background for his own views. In a sense it might be regarded as an *apologia pro sua vita,* at least for himself as a psychologist. He gives a rounded statement of his own position on most points without repeating the detailed evidence of his earlier works. He asserts that his plan is to "outline how far psychology has really made progress. . . . In particular we will see how far it has gone beyond the confines of common sense." It makes no claims to being a history of psychology. The other systems are brought in by way of allusion or by brief quotations.

The earlier chapters are devoted to saying what psychology is not. Chapter I points out that no one has been able to give a satisfactory definition of psychology and that such as have been offered do not agree among themselves. This, the author asserts, is not fatal, for many subjects are agreed upon in practice without formal agreement as to how the topic or discipline may be formulated. It is next asserted that formal agreement about physiological, especially cortical, processes is not necessary for the advance of psychology. Awareness of a general connection of the two with hope that more may be known some day permits psychology to go its own way. Philosophy, too, offers problems that are of interest to psychology, but psychology need not wait for their solution, still far in the future, to undertake its own task.

His discussion of method is limited to criticizing both the excessive analytic method and its most severe critics, in advocacy of a more eclectic and catholic approach. He would apparently welcome both analysis and synthesis, and any other effective means of determining truth. In Part B are six chapters devoted to an enumeration of claimants to be mental faculties. Spearman takes a relatively favorable attitude towards the possibility of dividing mental capacities into units. He considers intellect, intelligence, attention, sensory faculties, memory and imagination, and the 'orectic' faculties in separate chapters with the result that each is rejected with varying degrees of emphasis. Intelligence is too ambiguous in application and tests assumed to measure it are entirely inadequate. Attention suffers from differences in definition, but is consistent in application. The orectic group, emotions and appetites, are not sufficiently organized to be useful. In the last chapter of this part "Faculties or Chaos" Spearman shows a tendency to accept something like faculties, for his factors would be at least first cousins to faculties, but he dismisses all previous attempts to formulate particular ones as inadequate.

Then follow in Section C chapters on different concrete problems of psychology in somewhat the traditional order. Sensations and their aspects or attributes are considered with the outcome that the process of analysis is retained as a valid method in psychology. Relations are said to be real constituents of experience that have been largely overlooked. When rightly regarded they are approximately identical with forms. Perception is asserted to contain supplements to sensations. These are in part reproduced experiences but an inclination is shown to bring in something that approaches the super-sensible to explain the whole of the interpretation. A chapter

on the "Way of Regarding" ends with the statement, which he asserts to be new, 'that we distinguish between facts and the ways of observing them.' This is the outcome of a consideration of the phenomena previously discussed under the term attention, selection, grouping, and related processes. Thought, Mental Dynamics, Mental States, Complexes of Behavior, Unconscious Mind, I and the 'Self,' and Mental Unity receive separate chapters. These must necessarily be brief and are treated by running through the different approaches current at present and relating them to early theories. In most he comes to the conclusion that in modern discussion all is confusion and the Gordian knot can be cut by going back to common sense from which psychology started and accepting its dicta. For instance in discussing the self he runs through the definitions of the different types of selves from Aristotle to Calkins and finds in them no agreement and enlightenment, then turns to the experiments of Michotte and his pupils and finds evidence in them that occasionally there is a fleeting revelation of an active self. This confirms common sense.

The final chapter in Volume I is entitled "The Confusion that is Gestalt Psychology." The title is the same as Spearman's article in the Jubilee Volume of this JOURNAL. The spirit and general conclusions are the same but most of the material is different. The tone of the article is bitterly hostile to the School, so bitter that it probably will carry less weight than a better tempered discussion would have done. Every premiss of the group is questioned. Possibly the most characteristic statement is that they have failed in their experimental work as in their theory to distinguish clearly between the subjective and objective aspects of experience. One is the matter of experience, the other a way of regarding it. Spearman's assertion is that "one is objective in the sense of dealing with matter that can be true or false, and so can serve to guide behavior. Whereas the other, and subjective configuration, is nothing more than a way of grouping. It is not concerned with the matter regarded, but only with the manner of regarding."

Volume II gives an analysis of the psychological laws. After questioning the possibility of laws through two chapters he believes that he has shown that psychological laws exist even if facts do not quite conform to them. The test of a law is that it shall afford the power of foreseeing the future as well as of divining the past. He devotes a chapter each to the Association of Ideas, the Law of Retentivity, divided into Dispositions and Inertia, the Law of Constant Output, the Law of Fatigue, the Laws of Noegenesis or of the origin of knowledge, the alleged Laws of Orexis, the Laws of Basal Conditions, age, sex, etc. These chapters outline and on the whole defend the laws enumerated in the chapter headings. They give in outline the basic principles upon which all psychology must be built.

The final part, Part E, will undoubtedly prove of most interest to the reader and will have the most lasting value since it gives an outline statement of the more important laws that Spearmen has established in the investigations of himself and his followers. The first two chapters consist in arguments which dispose of all claimants to special faculties and all individual types. The question of the faculties is reopened, but with no new evidence adduced in their favor. All the more recent attempts to classify individuals are evaluated and all are found wanting. Especially they "all fail to show what goes with what, which must be the minimum that any system of classifying individuals can offer."

Chapters 39 to 42 give the more detailed statement of the Spearman methods

and conclusion. First is a defense of the correlation method as the only method that provides facts for the discussion of types or capacities. The older men are said to have been rich in imagination but very slightly critical. The modern critics of correlation methods are said to accuse it of treating unimportant phenomena. Spearman denies this and asserts that if true it would be an objection only to the application, not to the principle itself.

Chapter 40 on the "Discovery of *G*" is a brief exposition of the methods and results and a long statement of the early objections and the ways by which they have been overcome. The statement is made that the existence of *G* is universally admitted even by the earlier critics. The following chapter on the specific factors is a little more tentative. There is definite assertion that special abilities must be recognized which supplement *G* by explaining the different degrees of success of individuals in limited tasks. Some are broader some narrower in application. Verbal ability is said to have the greatest evidence for their existence. Scientific ability, and practical ability also have claims. Others are suggested as possible. The chapter on Orectic Factors summarizes the results of Webb's early study and of others in the same field. Here the *W* or will factor is placed alongside of *G*. Several more special factors are mentioned but with less certainty. A discussion of method considers factor analysis as presented by Thurstone and Kelly, but concludes that the older tetrad method is more reliable and quite as easy of application as the more recent claimants.

The book represents a notable reaction of a leading psychologist upon the problems of psychology. It is highly individual and lays claim to nothing else. It presupposes a wide knowledge of psychology to appreciate it as much of it is made up of references rather than of exposition. One who does not know the literature would need to read most of the references given to appreciate the work that has been done and the conclusions that are drawn. The psychologist will find it stimulating and for him it will constitute a convenient source for the latest position of the author on the problems that he has made peculiarly his own.

University of Michigan W. B. PILLSBURY

*Heilkunde und Volkstum auf Bali.* By WOLFGANG WECK. Stuttgart, Ferdinand Enke, 1937. Pp. xii, 248; 27 plates.

In "Heilkunde und Volkstum auf Bali" (Medicine and Folkdom of Bali), the author has intentionally disregarded the as yet inchoate scientific cure of human ills; for modern Balinese physicians, despite their occidental training, generally still adhere to the old traditional practices surrounded with magic and mysticism.

Dr. Weck, who had been for many years a government physician in the Dutch-Indian service (p. vi), presents in his book the results of his daily observations and ceaseless investigations, supported and enhanced by a personal study of the old, profuse, often almost illegible, Balinese literature (Lontars) which still serves as a physician's manual. The titles of 256 Balinese Lontars are given on pp. ix-xi. A medical Lontar contains usually a usada (precepts, recipes, remedies) and a tutor (doctrines, commentaries); it is written on leaves of the Rontal palm (p. 3) and has been rewritten from time to time. Assisted by reliable informants and native-bred practitioners, he was well equipped to trace the developments of Balinese medicine and to analyze its intrinsic relation to Bali's entire cultural background. To reveal the latter is the specific purpose of his work. Dealing with very heterogeneous

features, *i.e. Balyans* (practitioners), anatomy, physiology, obstetrics; cholera, leprosy, pemali, small-pox, crops, diagnosis, remedies and poisons, the author not only conserves skillfully and interestingly the organic interrelation of every day Balinese critical situations and cultural demeanor but also shows how their entire existence, as well as their medicine, is rooted in and ruled by the self-same early principles of magic and mystic speculations.

Remarkable is the predominant influence that Hinduism (first invasion in 5 B.C.) has had upon the original Balinese conceptions in furthering the unique cultural structure to which the magic charm of "this most beautiful island in the world" owes its ever young force of attraction. Indian mysticism is responsible for the religious rites that affect the minutest part of Balinese life. The sacred writings are of godly origin (pp. 8, 9); divine revelations in hypnosis (p. 242), auto-suggestion (p. 30), trances (pp. 13, 18), and dreams, the messages from Sanghyan Suksma (divine purity, p. 52), are momentous in their medicine. Benevolent gods and intriguing demons omnipresently bestow gifts, and disperse or bring evil. Rigid ritual precepts govern worship, work and play, birth, youth and marriage (pp. 99-125). Strict esthetic-hygenic functions, observed with prayers and moral rectitude, are proper preparation for sexual intercourse (pp. 103-105) in which the psychic qualities of the participants are of more importance than the physical, to themselves and their offspring (p. 106). In contrast to the lax morals of other primitives, virginity is regarded as a virtue, a requisite for marriage; wedlock among blood-relatives is forbidden; matrimony's purpose is propagation (p. 108). Immoderateness in diet is sinful and punished by the gods (pp. 98-99). Thus, does religion find its realization in the behavior of this joyous people, living in a well coördinated form of agrarian socialism.

Rigid individualism is despised; there is only collective individuality in the social functions and arts of the Balinese. Life is conceived of as a collective unity, causing personal misdeeds to reflect upon the community; demons wreak mischief upon children for their ancestor's sins (p. 137). Ostracism is the lot of parents to whom twins of different sex are born for their surmised prenatal incestuous activities in the mother's womb (p. 106). While polygamy, especially by members of higher castes, (p. 14) is indulged in (p. 126), illicit sexual intercourse is disdained (p. 205); incest is sinful and weakens the village; adultery is punished with reincarnation into a reptile (p. 95). Divorce is ethical only for weighty reasons (p. 100).

Undoubtedly Indian elements have encroached upon the Balinese conception that the genesis of human life and soul is engendered by an ancestral soul (p. 109). The doctrine of "Ruwa Bineda" (pp. 39-52), *i.e.* the two adversarious principles and the belief in the magic power of the "Kanda Empat" (fetal membranes, pp. 52-62), is of interest. Closely related to "Paracelsus" (about 1500 A.D.) teachings, is the Balinese belief that a sort of fluid, called Mrta, is the life nectar of the entire world, the dewas and all creatures (p. 50).

Man, endowed with reason, *i.e.* the ability to discern good from bad, to bring joy or misery to others, is considered the crown of creation (p. 54); he is an image (microcosmos, pp. 237-244) of the great world and is in unison with it; his essence is the same as that of the macrocosmos (p. 64). Although the moral quality of the ancestral soul and parental conduct influence the fate of the newborn (p. 109), there is still opportunity for individual development. Demons and human maledic-

tions (desti, pp. 182-204) cannot harm one (p. 186) whose thought and speech are righteous and whose conduct is according to ritual. Essential is the veneration of Ciwa (creator, enlightener, pp. 9, 74) through prayer, asceticism and yoga (pp. 231-237).

Soul imprisoned by the body, as a bird in a cage (pp. 41, 109), is not a separated entity, but forms the entire psychological behavior, based upon atma (spirit-emotions), budi (wisdom) and manah (thought). Manah can be good or bad; directed by budi, it determines the character. Character is identical with conduct (expressions) and is reflected in the eyes (p. 44). The male is superior to the female (generally symbolizing evil, pp. 132, 192), for, dominated by her feelings (passions), she usually lacks budi, and follows less logical reasoning. A man, mastered by passions, is a weakling, a failure in life (p. 203). Such passions, as also the "cakta" (drives, trends), betray egotism and voluptuousness. Good qualities are essentially altruistic (p. 96).

Sense-psychology and physiology, intermingled in mystical speculations, find explicit consideration in Balian literature (p. 92). Ten sense-organs are outlined; five cognitive (ears, skin, eyes, mouth, nose) and five motoric (mouth, legs, arms, genitalia, anus) (pp. 94-95).

The analysis of the act of perception reminds the reviewer somewhat of the quadriphasic passage of each psychological experience in the Ewald-Kretschner theory; for the genesis of perception involves four powers: manas (intention, will), angkara (consciousness), budi (knowledge) and cita (comprehension); manas lets the angkara issue by way of the eyes; budi recognizes on the basis of experience; and cita utilizes the perceived object. Since perception is a consciously willed procedure, it is sinful to perceive evil things. This shows clearly a voluntaristic ethical aspect (pp. 43, 95, 233), and would offer an interesting object for comparison with the voluntaristic epistemologies of Anselm of Canterbury, Martin Deutinger, and Heinrich Rickert.

The Balinese theory of perception, as do all physio-psychological conceptions, lacks systematic unity. The various concepts and terms are frequently interchanged according to time, locality and caste (p. 95), just as there is not found a homogeneousness of religious conceptions and ceremonies in the communities (p. 5); for, "nothing can be generalized in Bali" (p. 171). Only from a perspective can this heterogeneousness present a compatible unified structure.

From this viewpoint the Balinese, although imbued with the religious mysticism of India and engaging daily in its magic practices, exhibits physically and psychologically a healthy personality that has been strong enough to withstand the inroad of all alien influences. He worships an essential pacifism and portrays, after his own pattern, a mode of social life, in which the ultimate object is seemingly not "self" but the development of the individual upon moral principles in, through, and for the community. Through its emphasis of the world's reality and life-affirmation, the Balinese conception surpasses the basically ego-sacrificing, collective spiritual monism of old (Hinduism), modified (Paracelsus), and modern (Tagore) Indian thought; thus finding a joyful and proper expression in personal ethic activity for the benefit of all.

Dr. Weck's "Heilkunde und Volkstum auf Bali" differs from other publications recently published by enthusiastic visitors to this "enchanted isle of the South Seas"

in aim, scope, and method. Written in an apposite style and fecund with descriptions, it familiarizes the reader with local customs and a rare romance of tradition which tend toward the unique beauty of Bali. It reveals to the physician and the ethnopsychologist, as well as to the learned layman, many interesting yet unknown features of the daily life of this greediless and happy folk, who, in their primitivity, have for the modern civilized man, in many ways, the indicative, ever timely message: that ethical development of the individual is the only basis for a healthy social structure and peace amongst men.

New York LUDWIG SCHOPP

*A History of Psychology in Autobiography: Vol. III.* By JAMES ROWLAND ANGELL and others. Edited by Carl Murchison. Worcester, Mass. Clark University Press, 1936. Pp. xviii, 327.

This collection of psychologists' autobiographies is the last of the three volumes contemplated in 1928 for publication in 1930 (p. x). Due to circumstances not explained, its appearance has been long delayed—longer even than the publication-date indicates, for the volume was not issued until late in the summer of 1937. Volumes I and II, dated 1930 and 1932 respectively, each contains 15 autobiographical sketches. The first volume extends to 516 pages and the second to 407 pages. The present volume, with 13 autobiographies—James Rowland Angell, Frederic Charles Bartlett, Madison Bentley, Harvey A. Carr, Sante De Sanctis, Joseph Fröbes, Otto Klemm, Karl Marbe, Charles Samuel Myers, E. W. Scripture, Edward Lee Thorndike, John Broadus Watson, and Wilhelm Wirth—runs only to 327 pages.

Nearly a decade has passed since these articles were solicited for publication (see p. 63) and only one author (Bentley, pp. 63-67) appended supplementary material to his sketch when the galley proofs were sent him during the early spring of 1937. With this single exception, the biographies do not extend to the date of publication. The attentive reader would doubtlessly discover this for himself, but as a matter of permanent record every article should have carried its date of writing.

As in the preceding volumes (see the review of Volume I, this JOURNAL, 46, 1934, 168 f.), the biographies differ greatly in style. Some are written chronologically, others start at the date of writing and work backwards, and still others start nowhere and end nowhere. Some are reportorial, some are reflective and philosophical. Some are straightforward life-stories, others are histories of the authors' intellectual development, and at least two are mere résumés and reviews of the author's books and papers.

Basic biographical data—such as full name, date and place of birth, parentage, and early training—are frequently omitted. Of the thirteen authors, three do not give their full names: two omit the first of their given names, and two use initials. Six do not give the year of birth; of the seven who do, only three give the month and day. Seven do not state the place of birth; ten do not name or mention their parents or provide any other genealogical information; and nine make no reference to their childhood or early training. If this information is desired it has to be sought elsewhere; but where should one turn for it if not to an autobiography!

The editor should have indicated, when he solicited the manuscripts, the kind of article desired. Had he done so, the biographies would not have varied so widely in treatment and in range, and the final product would have been of far greater value to psychology. We might not have escaped the success stories, but we would

have been spared the résumés and reviews that masquerade under the title of autobiography.

Four of the articles are translations from the German (Fröbes, Klemm, Marbe, and Wirth) and one from the Italian (De Sanctis). The translations on the whole are well done. As in the two earlier volumes, portraits of the authors are run in the frontispiece.

Though a fourth volume was planned and subsequent issues, at intervals of three or four years, were anticipated when the Series was begun, the editor announces (p. x) that the present volume closes the Series. Despite the shortcomings deplored above, psychologists are indebted to the Clark Press for making these volumes available. K. M. D.

*Fundamentals of General Psychology.* By J. E. DASHIELL. New York, Houghton Mifflin Co., 1937. Pp. xxiii, 655.

Psychology—presumably because of its youth—is not a stabilized entity, but an evolving one. The psychology of today is very different from what it was in the days of Wundt, not only in scope, but in its relative emphasis upon problems. To be up to date a textbook must have not only modern facts, but modern scope, modern emphasis and modern bias. Measured by these standards Dashiell's "Fundamentals of General Psychology" is truly representative of present psychology.

Let me give a concrete example. Titchener's textbook has a chapter entitled Affection. It stresses *affective value* of *simple stimuli* for the *average observer*, and interprets the material in terms of an affective element. There is no chapter on affection in Dashiell. In its stead are two sections of a chapter on motivation which stress (preferential) *attitudes* of *individuals* towards *everyday events and objects*.

As indicated by the preceding paragraph, the point of view of the book is behavioristic rather than introspective. This is to be expected when one considers that its predecessor was entitled "Fundamentals of Objective Psychology." But the Dashiell of 1937 is very different from the Dashiell of 1928. His behaviorism is now not only tolerant of introspective data, but willing to present such data without bothersome—and bothering—translation into behavioristic language. In his introductory remarks on sensory functions he writes: "Our primary inquiry will be to see how our hypothetical subject reacts to the various forms of stimulation which we bring to bear upon him and what differences in his actions are explainable by the changes of stimulation; but we shall now and then amplify this account by asking him, 'How does it 'feel' or appear?'" To illustrate this procedure: the section on visual sensitivity includes a paragraph entitled Introspective Analysis of Light—and no sarcasm is implied, nor apology made.

I have mentioned above that this book had a predecessor. It would be well to state briefly the relation of the two books. Much of the old book is present verbatim in the new one, but so is much else. The earlier book had 18 chapters totalling 571 pages of 37 lines; the present book has 22 chapters totalling 640 pages of 40 lines. The greatest difference, however, is in scope and point of view. The new book is not only bigger, but better and broader.

The approach of the new book is biological rather than historico-philosophical. "Homo sapiens," Dashiell early points out, "is moved by what are in essence the same forces both without and within as are the lower animal forms. and he ex-

presses his energies in the same general types of actions and action-tendencies." Throughout the book this view is apparent in the frequent use of data from animal and child behavior to illustrate or elucidate the functions of man. The book is not, however, a comparative psychology. The "normal, human adult" remains the focus.

Another major feature is the inclusion of much material involving description and theory at a molar rather than at a molecular level. In his introduction Dashiell states that "the ultra-simple and flinty concepts of the molecular behaviorism have been liberalized." They certainly have. A chapter on "Some Personal Problems in Motivation," after presenting an analysis by Lewin of certain crises in children, deals with the major problems of adjustment in terms of 'dynamic' concepts such as frustration, compensation and rationalization. The chapter on perception begins with Köhler and reactions to relations, and includes Wertheimer's principles of grouping. The chapter on the Fundamentals of Learning covers some of Tolman's work and devotes much space to Köhler's experiments on apes.

Let there be no misunderstanding, however, in connection with Dashiell's recognitions of 'molar' concepts. Köhler, Lewin, Tolman and Wertheimer are all in the experimental tradition. There is little tendency for Dashiell to stray into the more fanciful fields of molar description at the borderline between psychology and sociology or philosophy.

A further feature is presentation of subject-matter in such a manner as to interest the student. The topics which are meaningful to the neophyte—*e.g.* heredity and motivation—are dealt with before those which require a knowledge of psychology so that their importance can be rightly appreciated—*e.g.* sensation, the nervous system. The "naturally interesting" topics are also given more space, relative to their systematic importance, than the "dull" ones. Emotion gets nearly as many pages (30) as all of the sensory functions together (35). Social behavior, personality and conditions of efficiency each get an entire chapter. Finally, the book is profuse with concrete illustrations. Again and again a chapter or section starts with the description of an experiment or of a specific problem.

Life is short, art is long. Marked adequacy in some respects entails lesser adequacy in others. What price does Dashiell's book pay for its remarkable breadth, its true representation of modern psychology as a whole? The book shows an occasional lack of integration. We have noted that Dashiell admits introspective material in addition to behavioral data. What is the relation of these two types of evidence? The chapter on sensory functions is little changed from the equivalent chapter in the earlier text. It does include introspective material, but does so grudgingly. The chapter on emotions, on the other hand, reads at times as though it had been written by Titchener. From the literature on these two subjects one would expect just the opposite emphasis on introspective and behavioral material. I cannot escape the feeling that Dashiell's behaviorism and introspectionism need a bit more shaking to yield a good mixture. Another deficiency is a certain lack of depth and completeness in the treatment of specific fields. I do not find the startle pattern of Landis and Hunt in the chapter on emotion. Nothing is said about the "cochlear response" in connection with auditory sensitivity.

Do these deficiencies really weaken the book? Relative to perfection—yes. Relative to a practical standard—hardly at all. The book is not a theoretical treatise; it is not a handbook. It is a textbook for 'beginning courses.' As such, it is the best textbook that I know.

Harvard University J. G. BEEBE-CENTER

*Sleep Characteristics.* By N. KLEITMAN, F. J. MULLIN, N. R. COOPERMAN and S. TITELBAUM. Chicago, Chicago University Press, 1937. Pp. vi, 87.

Thirty-six experimental *Ss* slept for the authors in their usual surroundings for a total of 6800 subject-nights, 26 *Ss* serving at least 101 nights each, for an average of 241 nights. The authors sought for the effects on sleep of "a representative preparation of a 'food-drink,'" called *Ovaltine* or *Ovomaltine,* in doses of 8 grm. or 14 grm., dissolved in about 125 cc. of liquid. They wished to compare the influences of the solvents; namely, cold water, hot water, cold milk, and hot milk, on the effects of the solute, and compare the latter with the effects of the solvents taken alone. They also introduced for comparison the food equivalent of one sandwich or two sandwiches of bread and meat; and Amytal in doses of 1.5 grains and Evipal in doses of 4 gr. each.

The first characteristic of sleep which the authors studied belongs to a class of phenomena called "motility." While *S* was trying to sleep, and only at such times, his bedspring was connected to the pawl of a ratchet-wheel equipped with a rotation-counter. Whenever *S* moved from one gross bodily position into another, he advanced this wheel by one or more teeth. The time was measured in hours during which the apparatus was connected for operation. The authors counted the total number of teeth by which the wheel had advanced in this time, and divided it by this time expressed in hours. The result is "teeth per hour," hereafter denoted by the symbol TPH. To a few beds they also attached a clock which ran while the *Ss* were actually stirring and only then. From the paired readings of clock and counter they determined that the advancement of the wheel by one tooth corresponded roughly to 0.6 sec. of time spent in activity. They furthermore attached kymographs to a few beds to record the individual movements of *S* and their temporal separations. They did not present or describe the kymograph-records. I regard the omission as unfortunate.

According to their tables, the average *S* advanced the wheel through nearly 50 teeth per hour. By their rule this corresponds to about 30 sec. spent in activity per hour, which adds up to about 4 min. of activity in 8 hr. The activity is obviously lighter than walking on level ground. The total work-output that is so involved is absolutely small; it should make no great difference to a healthy person if it were quartered or quadrupled. The differences which they reported were usually smaller than 10%, amounting to about 3 sec. in an hour and 24 sec. in 8 hr. I consider them trivial in themselves, and I have no means of interpreting them as signs of some changes that may be important.

On the other hand, I feel that the kymographic records can be interpreted. If we can assume that while an *S* is lying still, he is disregarding some irritating bodily conditions that start when he assumes any bodily position, and increase until he stirs and relieves them, we may properly say that he is then inattentive to those conditions, unconscious of them, asleep to them, preoccupied with something else. But we then wish to know when he was lying still, and when he punctuated his immotility with movements, so that we can tell how long he kept quiet in each resting-pose, and in the average.

This information adds a great deal to records of total activity, and may forbid certain interpretations of them. For example, I once pointed out (*Psychol. Bull.,* 27, 1930, 1-39) that Van Leeuwen's dog, kept in a suspended cage connected to an apparatus like this one, turned the wheel through more revolutions on a certain

night when he had been overdosed with coffee than on another night after he had drunk Kaffee-Hag. Hence, Van Leeuwen reasoned, the coffee was very bad for the dog. But for these two nights, Van Leeuwen presented the kymograph records. They showed that the dog's extra evening activity after he had taken coffee occurred early; and that between 9:45 P.M. and 6 A.M. the animal was more quiet than he was in the same period of the Kaffee-Hag night.

Again, the present authors report that the mean TPH was not affected by atmospheric pressure, outside temperature, evening activity, season, or duration of sleep, and I believe them. But the kymograph-records of the activity of my *Ss* at Mellon Institute were affected by the last three of these variables. For some *Ss* the average rest-period is correlated more highly than 0.8 with the duration of their sleeping activity, this being measured from the beginning of the first rest-period longer than 7.5 min. to the end of the last rest-period shorter than 5 min. Hence, I do not feel that these authors chose the most sensitive means of detecting the effects of their experimental agents when they chose the mean TPH.

The group-average TPH for 20 *Ss* on nights when they took 14 grm. of Ovaltine in cold water was 44.9; on a *selected* group of control nights their mean TPH=53.6. But the mean TPH in 11 *other* groups of control nights was 47.2, instead of 53.6; it exceeds the experimental mean of 44.9 by only 2.3, instead of 8.7 TPH. The smaller difference would not be statistically significant; the larger difference is; but most of it is due to the *selection of control nights*. A similar accident accounts for a good part of the difference between the mean TPH following cold water containing 14 grm. Ovaltine versus cold water taken alone. But in 100 subject-experiments, 70 yielded a smaller TPH, 28 a larger TPH, and 2 the same TPH for Ovaltine in *some* solvent in comparison with control nights. The difference between the two sets of group-averages is about 8%; it is statistically reliable; and according to the authors' rule it represents a saving of about 2.2 sec. of activity per hour, or about 17 sec. in a night of 8 hr. (pp. 40 f.). In their discussion (p. 37) they treat this finding conservatively; but under the hands of the advertising copy-writers it behaved like the five barley loaves and two small fishes.

The *Ss* recorded the times when they turned out their bedside-lamps to go to sleep, and when they woke up they recorded also whether they slept intermittently or continuously; whether they dreamed or not, and if so, of what; and whether they felt well rested or not when they first awoke. In this set of criteria only the report "well rested" showed any effects of the substances ingested. In 59 subject-comparisons this report was associated with 14 grm. of Ovaltine more often than with the control; in 40, less often; and in 2 comparisons, equally often. The odds are about 19 to 1 that the difference was not due to chance. The authors admit that they required their *Ss* to pass judgments on what happened while they were drowsy, and therefore not fully competent to remember; they also admit that *S* might judge his quality of rest differently at different times after he woke up. The advertising-writers expressed no doubt of the validity of their interpretation; namely, that these results were highly favorable to Ovaltine.

From the *Ss*' records the authors derived what they call the durations of their sleep. The records indicated that if a person was awakened by external changes (such as his alarm-clock) he tended to sleep through a shorter time than if he was not so disturbed. I don't see why this experiment was necessary.

On p. 85 the authors assert: "Ovaltine in hot and cold water, or hot and cold milk, is the only material which, when taken prior to going to bed, increases the percentage of mornings on which one awakens well rested." For saying this they have been harshly criticized. It seems to me, however, that this should be taken for a slip of the pen, for in the preceding text they mention the effects of no materials except those which they tested.

National Research Council H. M. JOHNSON

*In the Realm of Mind.* By CHARLES S. MYERS. Cambridge, University Press; New York, Macmillan Company, 1937. Pp. 251.

Dr. Myers has here collected some of his lectures delivered between the years 1929-1935 before various business, educational, and scientific organizations. The subtitle reads: "Nine chapters on the applications and implications of psychology."

In Chapters I, II, and IV "The Help of Psychology in the Choice of a Career," "The Human Factor in Accidents," "A Psychological Regard of Medical Education," Dr. Myers crusades for the more extensive application of psychological tools and instruments, documenting his case with felicitous illustrations from the literature of applied psychology, and with illuminating comments on the rôle which psychology can play, *e.g.* in the selection, segregation, and training of medical students. The discussions contain much factual information, obviously written for the lay public.

Other chapters—The Psychology of Musical Appreciation, The Modern Development of Social Psychology, Towards Internationalism, and Psychological Conceptions in Other Sciences—provide challenging mixtures of fact, speculation, and opinion.

The last two chapters are on The Absurdity of Any Mind-Body Relation, and The Nature of Mind. In the first of these chapters, Dr. Myers repudiates all theories —psychophysical parallelism, interactionism, etc.—which seek to find a relation between mind and body. There is no relation to be sought, since Dr. Myers considers mind and body to be identical. The identity of "conscious processes and the processes of living brain matter" is established by a doctrine (in agreement with Hobhouse) which postulates two principles—direction and mechanism—as governing all activity, both living and lifeless. The two principles are inseparable in reality. Mechanical activities are controlled and guided by directive activities; directive activities are always involved in mechanical activities. "The characteristics of mind and life, in general, are identical. There can . . . be no possible *relation* between mind and living body."

In the last chapter, Dr. Myers insists that the subject matter of psychology is "conscious experience;" that behaviorists, while justified in claiming that we can measure only behavior, have no justification for their rejection of conscious experience as material for systematic observation and experiment. Conscious experience is inherent in "self-activity," which is simply a highly "distilled" and specialized form of universal directive activity. In the course of its evolution, consciousness became differentiated into (a) conative acts and processes of the self—striving, apprehending, etc.; (b) modifications of the self—the affects; and (c) presentations to the self—sensations, percepts, ideas, etc.

The book suffers from two common faults of collections of lectures. There is much duplication of material; and there is a good deal of incidental material. It is

impossible to indicate the wealth of discussion related to ethics, esthetics, the problem of measurement in psychology, Fechner's law, and Spearman's 'g.' The book is written in an easy and informal style.

Cornell University JOHN IRVING LACEY

*Feeling and Emotion: A History of Theories.* By H. M. GARDINER, R. C. METCALF, and J. G. BEEBE-CENTER. New York, American Book Co., 1937. Pp. xiv, 445.

The book contains eleven chapters dealing with feeling and emotion from the historical point of view. The material in the first three chapters was published separately by the late Professor Gardiner before his retirement from Smith College. After retirement he devoted a large part of his time to the completion of the present work, carrying the historical account from the Greeks to the end of the eighteenth century. He finished the first eight chapters and part of the ninth. These eight chapters were revised and the ninth completed by Dr. Ruth Clark Metcalf, his pupil and former colleague. The two final chapters, dealing with the development of affective psychology in the nineteenth and twentieth centuries, were written by Dr. John G. Beebe-Center.

A survey of the ample index and bibliography will explain the embarrassment of the reviewer about knowing how best to summarize the content of the book. There are so many names and theories, so much rich detail, careful consideration and appraisal of views that the reviewer's task is not unlike that of describing an abridged encyclopedia.

It is more to the point to indicate the value of this book to the psychologist and the way it may be used with profit. If a student is interested in the history of psychology, and if he is especially eager to survey the whole sweep of affective psychology down through the ages, this is a book to which he can turn with the greatest profit. If eager to learn the views upon affection of Plotinus, Augustine, Aquinas, Descartes, Hume, Darwin, or someone else, he would do well to turn to the pages and bibliography of this book.

Apart from the value of the work for historical survey and reference, the last chapter is splendid to orient the student in contemporary theories of feeling and emotion. The reviewer suggests that in a college course upon feeling and emotion the last chapter could be profitably assigned to give perspective and orientation.

In the twentieth century the trend is away from theories which stress the primacy of consciousness and of behavior towards theories which recognize the primacy of the organism. Both consciousness and behavior are ephemeral and non-causal; correlations of these processes with organic are not fundamental explanations in terms of cause and effect. Today the dominant theories of feeling and emotion are *organic* theories, which stress the causal relations within a real permanent organism. These organic processes offer an explanation of certain organic events, themselves feelings and emotions.

Contemporary theories of feeling and emotion are classified by Dr. Beebe-Center as follows: (1) *Psychophysiological theories* which set out from an elementaristic conception of pleasantness and unpleasantness, and which attempt a physiological interpretation either in terms of peripheral processes (Titchener, Warren, Nafe) or in terms of central activity (Troland, Marston). (2) *Objective theories,* based upon the view that objective psychology is the only scientific psychology, which aim

to explain emotional behavior as reflex or as a pattern-reaction (Bechterew, Watson). (3) *Organic theories* of two kinds. (a) *Dynamic organic theories* which emphasize the real, permanent organism and regard conscious and behavioral facts as merely indices of organic processes. The view is shared by a group of men with psychiatric training and outlook (Freud, Janet, MacCurdy, McDougall). For these men emotion is in some sense abnormal and the motivational factors are of basic importance in their interpretations. The work of Lewin and Luria has aimed at an experimental approach along the same line (b) *Physiological organic theories* are propounded by men of physiological training (Head, Cannon, Dumas). These men stress the explanation of affective organic processes in strictly bodily terms. In the last chapter is a very useful summary and appraisal of the work of Dumas, who is perhaps not so widely known among American psychologists as Cannon and Head.

Now as to impressions from the book. Down through the centuries there has been a practical, moral angle to the study of feeling and emotion. Pleasure-pain, the sentiments, emotions, passions, have been viewed in relation to practical motives, to virtue and vice, to the happiness and dignity of man. This note was sounded in the ancient Greek writings and on through the centuries. Eighteenth century discussions of egoism and altruism, of hedonism and standards of conduct, preserve the tradition. Present day emphasis upon application is not new.

The book is evidence for the statement that we have always had plenty of speculation about the passions, emotions, feelings, and similar topics. The historical scholar, willing to dig deeply into the past and bring up the skeletons there found, has proved the point. It is illuminating also to discover that apparently recent views—the James-Lange theory, modern hedonism, the judgmental theory of affection—can be duplicated in the past.

The thought is a bit depressing that after so many centuries of speculation about feeling and emotion we still have so little that is settled and stable today in this field. But the outlook is somewhat brighter. Perhaps the objective, experimental approach of modern times will show law and order where previous centuries of speculation have yielded only shifting verbal definition.

It is inevitable in a book of this kind that some of the views be treated sketchily and others more at length, because of the varying amounts of available material, the varying historical importance of different theories, and other considerations.

The book as a whole is carefully written and clearly printed. It is a credit to both authors and publisher and a permanent contribution to the history of psychology.

University of Illinois PAUL THOMAS YOUNG

*The Definition of Psychology.* By FRED S. KELLER. New York, D. Appleton-Century Co., 1937. Pp. vi, 111.

A dozen years ago psychologists, having come a certain distance along a tortuous path of theory, looked back upon their brief and unimposing history and took account of stock. What they found readily lent itself to cataloging and pigeonholing of points of view; the "schools" suddenly became an important and timely topic. Old distinctions were revived and new ones invented to insure that the various camps would be kept clearly separate. At least distinctions were clearly enough drawn to cause camp-followers no inconvenience in arriving at decisions as

to affiliations. It was not generally assumed, however, that family quarrels within psychology would assume such basic importance that one day the beginning student would find it necessary, before being subjected to an observable fact or two, to get himself oriented with respect to the "kinds" of psychology ("schools") available. Yet the author of the present book has this to say by way of preface: "This book is the partial attempt, during the past six years, to provide college students with an introductory course of instruction in psychology that will serve at once as an orientation for those who expect to go no further into the field and as a preview for those who plan to concentrate therein. . . . In my own teaching practice this material constitutes the first part of the course work. . . ."

It is a temptation to attempt an evaluation of this book by relating it to two earlier works in the Century Series, Heidbreder's *Seven Psychologies* and Boring's *A History of Experimental Psychology*. It is essentially a condensation of the two contents. Were it not for the extraordinary avowed purpose of the book, that of the orientation of undergraduates, one might properly question its right to serious attention; the historical task has been done so much better by Boring and the review of the "schools" so much more thoroughly by Miss Heidbreder. One may, however, disagree with the author as to the wisdom of introducing general psychology in this particular way and still concede that the book is very readable, even for undergraduates. It possesses many virtues in common with Woodworth's *Contemporary Schools* and is, in fact, more directly comparable with that book than any other single one.

The first two chapters lay the background for the discussion of modern viewpoints, the first presenting a brief sketch of the philosophical roots of speculative psychology and the second being devoted chiefly to the Wundtian influence. Successive chapters deal with "Titchener and Structuralism," "America and Functionalism," "Watson and Behaviorism," and "Germany and Gestalt." The final chapter is entitled "The Problem of Definition" and endeavors to show that through the operation of a variety of current trends (some of them questionable, such as "if the modern psychologist is less concerned with consciousness, he is also less concerned with physiology") psychologists are becoming eclectic.

Throughout the author displays an ability to get within the various systems and portray them sympathetically, though it cannot be said that all viewpoints are treated equally in this respect. The chapter on Gestalt psychology, for example, lacks both breadth and clarity. It is, in fact, a better review of Koffka's 1935 book than a general exposition of Gestalttheorie. A similar restriction of the discussion of "functionalism" to that of the Chicago school partially defeats the author's purpose.

Since the book makes no pretense to scholarly originality the psychologist should not expect to be informed by it. In fact, the reading of it will not be repaid unless through the discovery of an altogether charming literary style.

University of Virginia FRANK A. GELDARD

*Theory and Practice of Psychiatry.* By W. S. SADLER, St. Louis, C. V. Mosby Co., 1936. Pp. xxii, 1231.

The psychologist who reads only the introductory chapter of this book will almost inevitably damn it as a sorry affair. Those who go beyond the early sections to complete a critical reading of the 1200 pages of the volume—as this reviewer

avers he has done—will find it less easy to express an unequivocal opinion. It may be enough to remark that the content is as varied as the style of presentation which shifts without warning from the chatty breeziness of the popular lecturer to the serious and well-informed exposition of the physician addressing his colleagues.

The introductory section establishes a 'low' for the volume. It is carelessly written and suggests a superficial acquaintance limited to secondary sources. Here one learns that Wundt founded a school which "ever since has borne his name" and which produced researches that "were soon coordinated with those of the Russian physiologist, Pavlov." Here one is told of "the opposing schools of 'introspectionists' and 'structuralists' "—the former being composed of self-psycholgies. Here one learns—not without some amazement—that Titchener espoused a cause called "structuralist-behaviorism," although "he made peculiar admissions for an out-and-out structuralist when he alluded to his 'unity of experience'." Let us draw a curtain of charity over the chapter. It must be read *in toto* to be appreciated.

The first major section of the volume is labelled "Theory of Psychiatry;" it presents an eclectic treatment of the background upon which modern psychiatry rests. Here, as in other psychiatric treatises, one is impressed by the abundance of agents and causal entities. It is indeed a long step from current psychology and its search for a single agent to psychiatry and its ready multiplicity of agents—but here Sadler hews closer to the line than many of his contemporaries. The second section of the volume deals with "Personality Problems;" it involves a discussion of psychogenesis from infancy to adulthood and includes a novel and stimulating section on problems in the nursery and in the preschool period. The fourth group of chapters deals with the psychoses and leaves the reader with an unintended but none the less sharp realization of the primitive level of therapeutics in this large field. The concluding—and perhaps the best—section of the volume presents the author's approach to psychotherapeutics.

At his best the author is a keen clinical observer with an eye for accurate reporting of symptoms and an ear for recent developments in his field. At his worst he is an opinionated 'Doctor' who settles moot points by simple proclamation and who holds no brief for "all this modern palaver about scientific child culture." He seems unable to decide upon the audience to which he is writing. When he appears to have the layman in mind, he is pedantic and willing to solve any problem by creating an epigram. When he seems to address his colleagues, he can be careful, critical, and thought-provoking. When he writes of psychiatry as a curative art, he can write splendid sections on the pitfalls of diagnosis. When he attempts to don the robes of the philosopher or the scientist, he reveals himself as an unconscious master in the use of the logical fallacies, with special emphasis on the production of imposing *non sequiturs*.

If one is willing to ignore frequent self-contradictions and an inevitable dodging behind figurative expression whenever the way becomes difficult, and if one can forgive a literary style which smacks of the house-by-the-side-of-the-road tradition, one is left with a volume which presents a wealth of clinical evidence with frequent signs of originality in interpretation. The physical bulk of the volume will make it an imposing addition to the shelves of ambitious young physicians but at the same time is likely to confer on it the dignity of a standard reference work which ill befits so casual an exposition. Psychologists who are willing to read critically and

patiently will find its better parts a stimulating compendium of one man's clinical experience.

University of Maryland JOHN G. JENKINS

*Introduction to Clinical Psychology.* By EDWARD M. WESTBURGH. Philadelphia, P. Blakeston's Son and Company, Inc., 1937. Pp. xiii, 336.

The clinical aspects of psychology, which have been slowly developing, appear to be reaching a stage of adulthood in which they are being expounded in a number of books. Dr. Westburgh writes an introduction which is a genuine contribution to the subject. He has continually in mind the student who already knows something of the academic approach. An attempt is made to give such a student a new viewpoint, in which clinical needs and attitudes are in the forefront. But, at the same time, the author definitely refrains from the temptation to write a mere manual of clinical procedures, which might be embellished and padded with numerous case histories. The omission of these adds to the readability of the book and keeps it on a higher plane.

The work opens with a brief historical discussion, in which much emphasis is given to the psychoanalytic school by means of a lengthy quotation from Freud. This is followed by a discussion of mental tests, in which it is assumed that the reader understands the nature of tests, so that the stress is placed upon their interpretation. The statistical meaning of test results is considered in great detail, starting with the most elementary ideas, and much is made of the use of percentiles as interpretive instruments. One notes with satisfaction the tendency away from the uncritical use of the intelligence quotient, and regrets an unfortunate mistake (p. 46) in the interpretation of the "probable error."

The next two chapters, devoted to "Cognitive Factors" are less satisfactory. They contain theoretical considerations, the result of the author's training with Witmer, in which a hierarchy of different kinds of intelligence is elaborated in some detail. The reader tends to be confused by the maze into which he is led. The practically minded person questions the utility of the discussion, and the theorist feels doubts as to its scientific validity. After a consideration of the developmental and intellectual history, there come two chapters on "Affective Factors" which express vastly different points of view. The first is essentially an exposition of McDougall's ideas on feelings, emotions and sentiments. The second, however, is concerned with an elaboration of the mental mechanisms which have been brought to light by psychoanalysis and its offshoot schools of thought. These mechanisms are explained with unusual clarity, and are nicely related to the theoretical considerations of the preceding chapter. The text now returns to further directions as to the taking of the patient's history in its social, vocational, affective and physical aspects. The final chapter deals in some detail with the various psychotherapeutic procedures available to the clinician.

Not only must the purpose of this book be kept in mind when evaluating it, but also the author's concept of clinical psychology. His insistence that it is something more than psychometric testing is not only commendable, but vitally necessary in this day and age. But the field covered in this book is large, and may, perhaps, set too ambitious a goal for the clinical psychologist. One who would follow Dr. Westburgh's ideas would perform all of the examinations (except the physical) included

in a complete study in one of our modern mental hygiene or child guidance clinics. But such a clinic finds it necessary to have specialization among workers, and divides the task among psychiatric social worker, clinical psychologist, and psychiatrist. The single clinician who can cover the entire field with only psychological training must be a rather exceptional person—or else must do an inadequate piece of work.

These considerations, however, must not be allowed to detract from the value of a very clear and scholarly presentation of the psychological problems, a presentation that is admirably suited to the purpose for which it is written.

Milwaukee County Mental Hygiene Clinic GILBERT J. RICH

*Social Psychology.* By JAMES MELVIN REINHARDT. New York, J. B. Lippincott Co., 1938, Pp. ix, 467.

In the central theme of this book, the socio-economic or cultural determination of personality, Professor Reinhardt joins the company of contemporary anthropologists and one or two social psychologists. More explicitely, he draws attention to the circumstance that the cultural conditions which mold personality resolve themselves into economic factors, and that the ideas, judgments, emotions—in brief, the values—of a group will be ultimately determined by the particular social relationships which issue from the economic interests at stake. Throughout the book, but emphatically in the last chapter, he deals with the 'rigidification' of institutions, formulas, fantasies, and frauds that are of good repute, by which the individual paradoxically acquiesces to his own loss. Once a practice is established and maintained by the interests of those who gain, or believe themselves to gain, from it, it becomes integral with personality by the sheer fact of its existence, carries its own convincing proof, and automatically ignores contradictions and irrationalities. Professor Reinhardt is very definite on this point as a generalization concerning all human groups and aboriginal groups in particular. Although he is not quite so explicit on the matter in regard to our society, still no doubt is left that he can endure the implications of the principle for civilized groups also. This becomes apparent from his timely and skillful discussion of the effects of insecurity, particularly during the depression, upon personality, and from his general insistence upon the influence exerted on personality by the economic conditions under which people must pass their lives.

On the whole, the orientation of this work is sociological rather than psychological, and of the recent works written by sociologists it is the social psychology of preference for the instruction of social workers. It is not clear whether the book is intended for those entirely innocent of psychology or whether it presupposes a prerequisite. If the former, then the earlier material on nature-nurture, central nervous system, viscera, and glands is quite suitable. But if the latter, it would seem somewhat disproportionately emphasized, for these topics will already have been extensively treated in general psychology. While nature-nurture, especially in consideration of the rather novel data upon which the discussion is based, might be justified from either view, there seems to be little more than a concession to tradition in the inclusion of the other less functional matters. After all, if so very little is accounted for in general psychology by the nervous system, much less will be accounted for in social psychology. While the field of social psychology is debatable on numerous points, it would seem that these more purely structural and physio-

logical matters could well have been neglected for the sake of other, socially more vital and germane, topics, which have the additional claim of being ignored elsewhere.

Among these latter is the question of the group-mind, which though disposed of for professional psychologists is far from dead for others. While an explicit refutation of this fallacy would seem like the flogging of a dead horse, the fact remains that it is most definitely a live issue to students and others who continue glibly to speak of 'the spirit of an age,' 'the genius of a people,' 'mob spirit,' and 'crowd mentality.' Since Ginsburg, of the London School of Economics, definitely disposed of Le Bon's fallacy and Floyd Allport followed with his *coup de grâce* in this country, it has been assumed by the writers of social psychologies that nothing more remained to be said on the question. Psychologists may be convinced, but to the inexpert, who still continue to be influenced by this persistent notion and to accept it by implication in common terminology, the danger of the fallacy still remains. What is more, it is today being exploited over a large portion of the world to win men over to arduous and hazardous undertakings. It is implied in the self-sacrifices which are demanded on behalf of the dubious ambiguities of racialism and nationalism. In such cases the concept of the group-mind is not taken as a metaphor symbolizing the common cultural heritage, but as a literal entity with a definite superorganic existence in some superindividual dimension. Since such a view is as likely to be found in the classroom as on the curbstone, however vague and unformulated it may be, it would seem wise to pay some attention to it. Teaching is as much a matter of disturbing misconceptions as it is a matter of instilling sound notions. Writers and teachers of biology do not weary of reiterating the, to them, tedious objection to the spontaneous generation of life, because there happen to be many students still very much in need of it, however obvious it may now be to the initiate. It should be acknowledged that, while Professor Reinhardt neither specifically treats, nor for that matter, names, this fallacy, he touches upon it tangentially in connection with the power of the cultural heritage of the group. But this is not quite enough for a misconception of such persistence, vitality, and especially contemporary significance. It ought to be named and unequivocally repudiated, as biologists repudiate the other.

The style of writing is compact, in fact sometimes too terse for elementary students. It is also cultivated and scholarly, without undue concessions to the limitations of undergraduates. These merits will probably cause inferior students to complain of it as difficult. But good students will find it stimulating and interesting.

Sarasota, Florida ELLIS FREEMAN

*Controlling Human Behavior.* By DANIEL STARCH, HAZEL M. STANTON, WILHELMINE KOERTH, assisted by ROGER A. BARTON. New York, Macmillan Co., 1936. Pp. xiv, 638.

This book, writes the senior author, "is designed to be a first book in psychology for use in introductory courses in colleges and universities. It comprises much material not found in current texts in addition to what I and the collaborating authors believe to be the essential scientific material for an introductory book" (p. vii).

A large amount of the material in this book is presented in a way to interest the

beginning student and several chapters, *e.g.*, Chapters 13 and 24, are in the reviewer's opinion excellently written indeed. The authors are successful in many areas of the book in focusing attention on matters of practical concern to the student. Generous bibliographies are offered and "Questions for Yourself" are found at the end of each chapter. For a number of reasons, however, the book is disappointing.

In the first place the book is ill-conceived. When a dictator, a highpressure salesman and one attempting to make love or to practive the Golden Rule are all said to be *controlling human behavior* the term becomes psychologically and descriptively meaningless. One finds here that *control* means anything from incarceration of "unregenerate criminals" (p. 483), to control of emotional patterns (p. 398), control of imaginative activities (p. 387), control of motor and mental development (p. 386), self-control, winning one's sweetheart or selling a hundred million dollars worth of automobiles (p. 22). The collaborating authors give the impression of having had to force the use of this word, though perhaps it was not force but some other kind of control.

In the second place, the book cannot maintain the pretentions of a scientific approach. The responsibility for this condition rests on the senior author. In his first two chapters, for example, his castigation of the Russian economic experiment and his extolling of systems giving free play to the "profit motive" are biased, unsupported and beside the point. The senior author also imputes validity and reliability to the psychological observations of individual newspaper reporters, economists, Vassar graduates as such, and fashion experts (pp. 9; 13; 30, three citations).

The book as a whole is permeated with loose statements and figurative expressions that are psychologically meaningless. No one of these expressions is especially important in itself, but the large number to be found will make it difficult for both instructor and student.

The senior author gives the appearance of being deterministic and of having the answers. "You will probably learn that one of the basic principles of human nature is not to force the other person except rarely, but rather to guide his thinking toward the goal you desire—toward the right goal. Therefore make sure that the goal you desire is the right goal for you as well as the other persons in the situation" (p. 22). Having assured oneself that he has the right goal, he gets what he wants by "psychologizing the other fellow" and the situation (p. 17).

The senior author emphasizes three kinds of control: "psychologizing" or getting the other person to think out his own problems in the "right way" (your way); the use of *power over other people* (pp. 28, 31, 32); and the use of the Golden Rule (pp. 597 ff).

The *power-over* "leader" described by the senior author in the beginning of the book is not the same as, nor is he differentiated from, the *integrative* "leader" discussed by one of the collaborating authors in Chapter 13.

A few bibliographical errors have been noted in passing. Woodworth's *Adjustment and Mastery* is reported as published by two different firms (pp. 371 and 487). The names of William Healy and of K. de Schweinitz are consistently misspelled throughout the book.

University of Illinois HAROLD H. ANDERSON

*Metodi Compiti e Limiti della Psicologia nello Studio e nella Prevenzione della Delinquenza.* By A. GEMELLI. Milano, Società editrice "Vita e Pensiero," 1936. Pp. xiv, 155.

In Italy the study of criminology has generally proceeded on anthropological and sociological lines, because of the impulse given by Lombroso and Ferri and the apparent successes of the positivist school which grew out of their efforts. This school, primarily concerned with studies of heredity, anthrometric measurements, and the discovery of biological anomalies or stigmata as supposed indices of criminality, has expanded to include studies of endocrinology and physical constitution, but still neglects psychology and its methods. There is consequently no Italian literature on criminal psychology comparable to that which exists in Germany and the United States.

It is partly to remedy this condition and partly for the purpose of subjecting the anthropological school to a searching criticism that Professor Gemelli has assembled this group of papers read at various conventions of Italian psychologists and anthropologists. These papers, as the title suggests, deal with the methods, functions and limitations of psychology in the study and prevention of delinquency, but include further an account of the merits as well as the limitations of criminal anthropology, and a psychological analysis of the new Italian penal code introduced by Minister Rocco. The positive school of criminology served a useful purpose in opposing the classical school of criminology which tended to barren speculations and an emphasis on legal abstractions. The positive school also led to the recognition of a considerable class of criminals and delinquents as cases of illness or mental defect and, therefore, as objects of medical care, rather than of penal discipline. But it erred, according to Professor Gemelli in trying to include all cases in this category, and in its assumptions of materialism and fatalism. Moreover, the positive school ignores the most important data from which delinquency can be determined, viz. the psychology of the individual.

Agreeing in general with De Sanctis concerning the function of psychology in the study of the delinquent, Professor Gemelli lays down certain methodological premises for the study of the delinquent, whereby a complete and satisfactory scientific approach may be attained. Anthropological study is necessary to account for the influence of organic factors on the delinquent, such as heredity, constitution and morphology. Sociology is equally necessary for an account of the influence of environmental factors. But psychology, as the study of the phenomenology of psychic activity, is absolutely essential in determining the mechanism of delinquent behavior. None of these sciences must claim to supply all of the relevant data, or proceed beyond its proper limits as a definite science, to the prescription of legal forms or procedures or the construction of a philosophy.

The psychology appropriate to the study of the delinquent must recognize will as a genuine psychological factor, a function of self determination in the presence of motives. To prove that such a function exists Professor Gemelli relies on studies growing out of the work of the Würzburg School. In his capacity of psychologist he refuses to invoke the philosophical assumptions of the Neo-Scholasticism to which he adheres as a philosopher. In so doing, the learned Rector of the Catholic University of the Sacred Heart is both scientific and indulgent toward opponents who could not follow him on metaphysical ground.

The special methods of psychology are of unequal value according to the author. Characterology is of little use in the study of the individual, although it may be of great theoretical value in the investigation of problems of general psychogenesis. Psychoanalysis presents an abstract classification of delinquents of no particular originality, but offers a valuable method of studying the individual in its technique for the exploration of the unconscious. The greatest contribution must come from differential psychology which is the only method capable of rendering precise and detailed, the general picture of the delinquent offered by the other methods and sciences.

In conclusion the author offers a provisional classification of delinquents in five categories determined by variations in affect, volition and habit, and the intensity of initiating stimuli. Delinquents, the author repeatedly insists, are normal people with psychological processes of almost inconceivable complexity, and the two tasks of psychology are to furnish the judge with the analysis of delinquent behavior, facilitating justice, and the penologist with a diagnosis which will be a guide to reform.

Some repetition is inevitable in a collection of independent essays, but this does not detract from the value of this interesting work, characterized by learning, an attitude humanitarian as well as scientific, and an unvarying courtesy toward representatives of other doctrines.

University of Cincinnati CHARLES M. DISERENS

*Psychology and the Social Order*, by J. F. BROWN. New York, McGraw-Hill Book Co., 1936. Pp. 500.

Social psychology has been for a long time more a battlefield between sociology and psychology, where each of these sciences defended its own view against the other, than a linkage between these sciences. Recently there are signs of a change to a more coöperative attitude and Brown's book is a decided step in that direction.

Although the history of sociology differs in some respects from that of psychology, certain methodological problems of today are strikingly similar, *e.g.* the intermingling of historical and a historical systematical questions, and the increasing need to develop dynamical, yet empirical explanations instead of mere description and classification. Brown stresses the necessity of a definite transition to a "field theoretical" approach. He discusses the mathematical means which have been developed in topological and vector psychology to represent psychological fields and tries to apply the field theoretical approach to the outstanding problems in social psychology. Groups are viewed as social fields. A scientific representation is offered of such factors as the degree of freedom of social locomotion, the unity of a group, the degree of separation of groups or sub-groups, membership in a group or in several overlapping groups.

With this methodology as a tool he attacks a field much larger than generally implied in social psychology. Marx, Lenin, and Spengler are as much discussed as Freud and McDougall; problems of rural and urban families, race and nationality, no less than the more traditional problems of social psychology such as instinct, group mind, social attitude and leadership. He applies the field theoretical view to a number of more specific problems, such as the effect of war on the structure of a nation and its classes. In the political science section of his book he approaches

the problem of the state and the different types of government, such as democracy, fascist and communistic dictatorship.

In presenting a new methodological attack of a general character on such a large field the scientist has to choose between two possibilities: he can either try to go first into detailed specific work, presenting the methodological implications only afterwards; or he can come out first with the presentation of the broad view, leaving the proof of fruitfulness to the following detailed studies. Both ways have their definite virtues and definite difficulties. Brown has chosen the second way. His discourse is often brilliant and intriguing. Sometimes the broad nature of his attempt has hindered a more thorough analysis and has led occasionally to distinctions which are more schematic than explanatory. The field theoretical approach makes it necessary, as Brown points out, to develop the conceptual side of the constructs together with elaborate investigation of the content. That of course is possible only in detailed studies but I am persuaded that the future of social psychology and of sociology will bear out the fruitfulness of the field theoretical approach.

University of Iowa KURT LEWIN

*The Psychology of Early Growth.* By ARNOLD GESELL and HELEN THOMPSON. New York, Macmillan Co., 1938. Pp. vii, 290.

This book is to be considered in a sense as an elaboration, within a restricted field, of Dr. Gesell's earlier work, *The Mental Growth of the Preschool Child,* and as a more complete and systematic report of the material presented in *An Atlas of Infant Behavior.* It describes the results of a normative study of infants between the ages of 4 and 56 weeks which has been in progress during the past 10 years. The main part of the book is an exposition of the procedure and results of this study. The presentation shows many admirable points. At no time is the reader in any doubt as to what was done. The choice of subjects and the characteristics of the normative group are discussed in detail. This group was essentially a homogeneous one, composed of healthy children of native-born parents of northern European extraction and of middle socio-economic status, rather than a random sample of the population. Such a choice accentuated the central trends of development which were the primary interest of the investigation. The experimental procedure is expounded with equal clarity and exactness. Some 30 types of situation were presented to the infants, with studies at every 4-week age-interval, and the behavior carefully recorded and classified. Anthropometic measures were also made.

The behavior norms are presented in a number of ways: raw tabulations of response frequencies; summaries of different types of behavior at each age-level; as a functional syllabus of the temporal development of every form of activity. From this data, a schedule of development, entitled an "Analytic Developmental Schedule," is elaborated. It is based upon the scoring of three types of behavior items: those which appear with increasing maturity; those which disappear in the process of growth; and those which first appear and then disappear, but reach maxima at definite focal age-levels. A method is given for computing approximate developmental ages, and separate schedules are provided for postural, prehensory, perceptual, adaptive and language-social behavior. The tendency to combine discrete observations into a single numerical figure is carefully guarded against.

The reviewer feels that the suggestions for the clinical use of the developmental schedules clearly imply that useful norms have now been set up. Is such an implication justified? The use of a homogeneous group of infants was ideal for the purpose of developmental study, when one age-level was to be compared to another. But one may question whether such a group is satisfactory as a norm to which infants not so chosen are to be compared. Until we know the effects of racial inheritance, socio-economic status of parents, etc., upon development at this age, errors are likely to result if an infant is compared to a standard which is alien to him. This is, of course, only a minor criticism of a rather monumental work, one which is outstanding because of its scientific thoroughness and the completeness with which every detail is presented.

Milwaukee County Mental Hygiene Clinic GILBERT J. RICH

*General Psychology: Briefer Course.* By FLOYD C. DOCKERAY. New York, Prentice-Hall, 1937. Pp. xix, 372.

This book is based on the author's *General Psychology* of 576 pages (reviewed in this JOURNAL, 48, 1936, 702) and is designed through omission of the more technical parts, for brief introductory courses lasting one quarter. This review will do no more than compare the two books and criticize some details which the former review did not.

The small book is derived from the large one by retaining in full 16 chapters and in more or less shortened or rearranged form the material of 7 others. The parts omitted are those on sensory discrimination (3 chapters), scientific method, the experiment, the organism, types of simple sensory-motor responses, and statistics. A good deal of cutting occurs in the chapters on learning; the material on principles of learning, with its references to animal experiments, is almost entirely left out, and the chapters on learning in the small book are headed "Practical Problems in Learning." The chapter on "Problem Solving in Apes and Men" is, however, retained in full. It appears that nothing was especially written for the small book except some illustrations "selected from concrete experiences of students in everyday situations."

The cutting down is carried very far in the sections on the nervous system (58 lines and 4 figures), vision (58 lines and 2 figures), and audition (33 lines and 3 figures). The reviewer doubts whether material so abbreviated retains any value at all. Fig. 9 is misleading, as the optic disk and the fovea are too close together. In Fig. 7, the structures of the inner ear are intentionally made too large in relation to other structures, but the student is not told this. The theory of hearing is given thus: "The vibrations are thence conducted by three small bones to the inner ear, the cochlea. Here tiny brain cells, the true receptors, are stimulated. It is probable that some respond to one frequency and others to other frequencies."

The chapter on emotion, because of omitted sections, makes no mention of the three divisions of the autonomic nervous system and says that "its chief characteristic is that it conducts impulses in a rather diffused manner to the internal organs." The large book, in fact, does not distinguish between the diffuse activity of the sympathetic segment and the specific innervation of the parasympathetic. It seems to the reviewer that the discussion of the conditioned reflex on p. 44 is very casual and that its importance is not emphasized as much as one would expect

from the author. On p. 273 we are mistakenly told that Terman "inaugurated the use of the intelligence quotient." Stern, as a matter of historical fact, originated it.

The style of the book is, as a rule, lively and good, but "none" is usually found with a plural verb. The pictures are few and generally unattractive; for example, the full-page picture of the chimpanzee on p. 20 is lacking in detail and out of focus, and the photograph of the multiple choice apparatus is not enlightening.

If the reviewer were giving a one-quarter course and wished to use a text by Dockeray, he would much prefer to have the students pay a little more and get the large text and to make his own or Dockeray's omissions. The process of omitting is easy, and in addition omitting always wins the applause of the students.

The Rice Institute FRANK A. PATTIE, JR.

*Introduction to the Rorschach Method.* By SAMUEL J. BECK. Monog. No. 1, American Orthopsychiatric Association, Menasha, Wisconsin, Banta Publishing Co., 1937. Pp. 278.

Dr. Beck has attempted to provide a comprehensive manual of Rorschach test procedure "to be helpful in the practical application of Rorschach tests." With this clear statement of the "sole" aim of this volume in the first chapter, together with the author's interesting declaration of insight into the artistic nature of that increasingly specialized and esoteric group of clinical workers known as "Rorschachers" (speaking as one without the fold), the further perusal of this book by a seriously interested but uninitiated outsider necessitates the learning of a new and difficult language of symbols with rather vague and slippery meanings attached to them and their combinations. Nevertheless, let us attempt to go further.

Characteristic responses to the Rorschach ink-blots with methods of scoring and interpretation are given for selected cases of healthy superior adults, the feeble-minded, depressed states, hypomanic conditions, schizophrenics, neurotics, conduct disorders in adults, and problem children. Fifty-nine cases are presented in all making up the major portion of the book. We cannot start here, however, for we are immediately confronted with a mass of special formulae, from whose mystical figures we find that the *O* is capable of a warmth of feeling, is sociable, of unstable mood, of very superior intelligence, and has a great deal of artistic sensitivity. *O* may be elated but will not have uncontrolled outbursts, is self-willed and has an intense inner life. The validity of this interpretation is not given. Some insight may be gained into the scoring methodology by turning first to the last 80 pages of the book (Part II, Technique, Norms, Interpretation) which should be read first in order to get a glimmer of understanding of Part I.

Some evidence as to the lack of validity in Rorschach interpretations is given in the chapter on "Inadequate Test Records" and evidence for a certain degree of validity is given in the differences shown between the cases selected as characteristic of certain clinical groups, but no statistical evaluation is attempted. Dr. Beck rightly assumes that this attempt at standardization of procedure must precede any serious attempt at the determination of either the reliability or validity of Rorschach methods.

One wonders if the construction of such an elaborate and involved methodological framework is justified by the quality of raw materials available from Rorschach responses. Also, there is a danger in the appearance of exactitude given by these numerical scores and formulae which tends to mask the fact that the scoring is still

essentially subjective. Beck fortunately does not adhere strictly to Rorschach's original typological classifications but attempts to obtain rough quantitative variables or directions of "processes or forces composing the personality." The processes or forces are essentially those suggested by Rorschach, *e.g.* Experience Balance (EB, Rorschach's Erlebnistypus), Energy converted into intellectual achievement (Z), Creative activity (M), Self-will (S), etc., with important modifications in their derivation and interpretation. The number of blanks and question marks left in the final interpretative table for the various personality groups studied are indicative of only a few of the many aspects of Rorschach procedure still in need of investigation.

Dr. Beck has made an important contribution in a most difficult field of investigation, one which frightens away the more (or less?) tough-minded investigator who would develop a neurosis of intellectual insecurity when faced with the problem of the total personality and its relation to society and had only ink-blots and Rorschach categories as technical equipment. Evaluation according to the usual scientific criteria seems pointless in a field which is and probably always will be reserved in its last analysis to the clinical artist and philosopher. The intelligent attempt to standardize a helpful adjunct to the psychiatric interview is commendable if the clinical values do not become swamped by a complicated methodological procedure.

Brown University HERBERT H. JASPER

*Principles of Educational Psychology.* By W. D. COMMINS. New York, Ronald Press Co., 1937. Pp. xiv, 596.

It is with some misgivings that the reader picks up *another* book on Educational Psychology in the presence of the flood of such texts in the last three or four years. The present text, however, is above the average book of this type both as to its difficulty in reading and in the amount of information it contains. It probably would be best suited for a class who had had a good basic course in general psychology.

The topics covered are as follows: Ability, General Intelligence, Learning, Transfer of Training, Motivation, Thinking, and Personality Development. The first chapter contains a brief résumé of the schools of psychology of most interest to the educator: Association, formal discipline, psychology of appreciation, physiological psychology, animal psychology, mental testing, and recent developments in the field.

Considering the class of students the book is obviously designed for and the number of topics brought together for consideration, the writer has managed to treat fairly clearly, accurately, and comprehensively the topics undertaken. Gestalt Psychology and the essential principles of Behaviorism have been given, it seems to me, a fairly well balanced emphasis in what is frankly an eclectic treatment in the field. The text has brought together a commendable amount of sound reading and should make an excellent working basis for a second year class in psychology preparing for teaching and later experimental work along this line. Graphs, charts, and tables are, all in all, presented clearly and in an understandable manner. Perhaps in a few cases a little more detail might have been desirable. The book is well worth trying out as a text with second year psychology students with an interest in education.

University of Pittsburgh W. T. ROOT

*Outline of Abnormal Psychology.* By RALEIGH M. DRAKE. New York, Longmans, Green and Co., 1937. Pp. 136.

This is one of the student outline series of which this is the third in psychology. The present volume can be divided into two parts, approximately equal in length. The first is a typical functional approach to abnormal psychology (pp. 11-73) in which sensory, perceptual, motor, associational, intellectual and similar functions are considered when behaving abnormally. The second part might be described as a disease entity or syndrome approach (pp. 74-107) in which the author discusses the psychoneuroses, the organic psychoses (in which the more common ones are included) and the functional psychoses. The remainder of the book is devoted to sleep and its abnormalities, dreams, hypnosis, suggestibility, psychotherapy and spiritistic phenomena. On the whole the volume can be said to be quite complete and has the very distinct advantage of including two approaches to abnormal psychology in one volume. It is rather unfortunate that the organization in an outline of this kind is not more uniform, so that when one is reading the material on theories of abnormalities of association (entitled incidentally, association theories) one does not have to encounter such arrangements as those to be found in this volume. The first five theories give the nature of the disorder and its cause, the next two discuss mechanisms (Freud) and forms of compensation (Adler) instead of causes. The causes are discussed under the topic of nature in these two. In this section, we also find a criticism of the psychoanalytic theories, but no criticism of the others. Similar deficiencies can be found in most chapters of the book. For instance, at the end of the chapter on functional psychoses there is a table describing the degree of integration, the kind of emotional expression, the personality type, the causes and characteristics of some of the functional disorders discussed in the preceding chapter. There is no such table after the chapters on the psychoneuroses or the organic psychoses. Incidentally, although the author lists the catatonic type of schizophrenia under functional disorders, he says its cause is "deficiency in nerve cell nutrients" in this table. Ignoring these somewhat minor indications of poor preparation of this work, it should be considered as an additional tool to be used in teaching students of abnormal psychology to learn definitions and lists of symptoms. So many of our students believe they are learning something about human nature in this way.

University of Newark FREDERICK J. GAUDET

*The Betty Book.* By STEWART EDWARD WHITE. New York, E. P. Dutton & Co., Inc. 1937. Pp. 302.

This book presents the philosophy of life that a young woman, Betty, gave in automatic writing and speaking. An appendix supplies some of Mr. White's observations of what happened in the eighteen years that these seances occurred. The greater part of the book is concerned with ethical and spiritual ideas which resemble a great deal of the material one gets from the better educated spiritualists. It has much to say about the difficulty of making contact with "the other side" and the necessity for more introversion among the pathetic extroverts. The appendix describes several mediums who have succeeded in making the so-called Beta Body visible. White admits that this material is not presented as scientific data. His interest centered upon obtaining the philosophy of the "Invisibles" rather than acquiring psychological data on automatisms. It is rather pathetic, however, to see a man like White tangled up in the sort of things that Betty spouts.

Johns Hopkins University H. C. MCCOMAS